ROBUSTNESS OF INCOMPATIBILITY FOR QUANTUM DEVICES

ERKKA HAAPASALO

ABSTRACT. A robustness measure for incompatibility of quantum devices in the lines of the robustness of entanglement is proposed. The concept of general robustness measures is first introduced in general convex-geometric settings and these ideas are then applied to measure how incompatible a given pair of quantum devices is. The robustness of quantum incompatibility is calculated in three special cases: a pair of Fourier-coupled rank-1 sharp observables, a pair of decodable channels, where decodability means left-invertibility by a channel, and a pair consisting of a rank-1 sharp observable and a decodable channel.

1. INTRODUCTION

As an inherently probabilistic construction, quantum theory abounds convex sets: the sets of states, observables, state changes, and measurements of a quantum system are all convex. Unlike in classical probability theories, these quantum theoretical convex structures are not simplexes, i.e., states and measurements cannot be decomposed into combinations of extreme points in a unique way. This gives rise to many of the interesting aspects of quantum theory.

The rich structure of the set of quantum states has been extensively studied; see [1] and references therein. Especially entanglement, as a truly quantum phenomenon, and its detection is the focus of great attention. In this paper, we concentrate on another peculiarity of quantum theory that has no counterpart in the classical world: incompatibility. Classical measurements can be carried out freely together and the measurements do not alter the system. On the quantum side, however, this no longer applies. There are many interesting pairs of quantum observables and measurements that do not allow any joint measurements or realizations. A canonical example is the position-momentum pair or any generalized Weyl pair.

In general, quantum incompatibility of a pair of quantum devices (observables, state-changes, instruments,...) is defined as the impossibility of joining the devices into a single quantum device from which the original devices could be obtained by reduction. We give rigorous definitions for incompatibility in all the cases studied in this paper but an all-encompassing definition of quantum incompatibility can be found, e.g., from [9]. It should be pointed out that the set of quantum states does not exhibit incompatibility; any pair of states can be joined into a bipartite state from which the original states can be obtained as partial traces.

Quantum incompatibility can be seen as a special resource like entanglement. That is, incompatibility is not simply a hindrance but rather a valuable non-classical feature that can be utilized in, e.g., quantum information processing. In fact, there are connections between entanglement and incompatibility: it was recently shown in [2, 8] that incompatibility of quantum observables and EPR-steering of quantum states are operationally linked. Incompatibility as a resource is thus strongly related to the resource theory of steering. Moreover, a quantum channel is entanglement braking if and only if its transpose maps any observable pair into a jointly measurable (compatible) pair [22].

There are several measures for quantum entanglement one of which is the *robustness of entanglement* originally presented in [24] that is purely based on the convex-geometric structure of the set of quantum states. Similar convexity based distance measures introduced for quantum convex sets include the *boundariness* defined in [10] and the *steerable weight* introduced in [23] quantifying the presence of EPR-steering. In this paper, we introduce a robustness measure for quantum incompatibility in the lines of robustness of entanglement. This quantity measures how well a given pair of quantum devices resists combining into a joint device under noise. Quantifying incompatibility of quantum observables has been earlier studied from a somewhat different viewpoint in [4, 12, 14], but here we extend the notion of robustness of incompatibility to encompass all relevant quantum measurement device pairs.

A general description of robustness measures is given in Section 2. In Section 3, we review the basic descriptions for the essential quantum apparati and, in Section 4, define the concept of compatibility of these apparati and introduce the robustness measures for incompatibility. In Section 4.2, we study some special properties of the robustness of incompatibility. We calculate the robustness of incompatibility in three exemplary cases in Section 5.

2. General robustness measures

The sets of measurement devices in any general statistical physical theory are naturally endowed with a convex structure. Namely, suppose that the set of devices under study is $\mathbf{Q}$ and Φ_1, $\Phi_2 \in \mathbf{Q}$ are devices of the same type. Then one can realize a device $\Phi \in \mathbf{Q}$ by applying Φ_1 with probability $t \in [0, 1]$ and by applying Φ_2 with probability $1 - t$. By expressing Φ as $t\Phi_1 + (1-t)\Phi_2$, $\mathbf{Q}$ becomes a convex set. In a convex combination $t\Phi_1 + (1-t)\Phi_2$, we view the coefficients t as random noise or perturbation caused by statistical mixing of devices, and we use the term 'noise' also in general convex geometries even in the absence of direct physical link. In what follows, we study convex sets K of devices or device pairs (which are also naturally convex by defining $t(\Phi_1, \Psi_1) + (1-t)(\Phi_2, \Psi_2) = (t\Phi_1 + (1-t)\Phi_2, t\Psi_1 + (1-t)\Psi_2))$ or, in general, *selection procedures* for a physical experiment with regards to a particular task or resource of the selections, such as entanglement (of individual states) or incompatibility (of device pairs) in the quantum case. We assume that there is a subset $L_0 \subset K$ useless selections with respect to the property we are studying and, moreover, that this set is convex as well, i.e., random mixtures of useless selections are also useless. In order to quantify the usefulness of an element $x \in K \setminus L_0$, we propose a way to measure the 'distance' of elements $x \in K$ from L_0.

Let V be a real vector space and $F \subset V$ an affine subspace. From now on, we fix a subset $L \subset F$ whose relative complement $F \setminus L =: L_0$ is convex; L_0 is to be thought of as the set of useless devices with respect to the task at hand. In the physical situations we will study later on, F will be the minimal affine subspace generated by both L and L_0, and L_0 will be *absorbing within* F in the sense that there is $y_0 \in L_0$ such that for any $x \in F$ there is $t \in (0, 1]$ such that $tx + (1-t)y_0 \in L_0$. These assumptions are, however, not necessary in this section. We measure the distance of $x \in F$ to L_0 by quantifying the least amount of noise (from L_0 or some other convex subset of F) to be added to an element $x \in F$ in order to enter L_0, or, in other words, how robustly x stays outside L_0 under added noise.

Let us make the following auxiliary definition.

Definition 1. For any x, $y \in F$, let us define

$$w_L(x|y) = \sup\{t \in [0, 1] \,|\, tx + (1 - t)y \in L_0\},$$

where we define $\sup \emptyset = 0$. We call $w_L(x|y)$ as the *relative L-robustness of x relative to y*.

It is immediate that, for a locally convex topological vector space V, whenever $y \in L_0$, we have $w_L(x|y) = 1$ if and only if x is in the closure of L_0. The number $1 - w_L(x|y)$ is

the least amount of noise in the form of a specific element y from F we have to add to x in order to enter L_0, i.e., the greatest additional noise in the form of y the element x tolerates without getting indiscriminable from L_0. It is immediate that, if V is a locally convex space and L_0 is closed, then the supremum in the definition of $w_L(x|y)$ is attained when $y \in L_0$, i.e., $w_L(x|y)x + (1 - w_L(x|y))y \in L_0$. See Figure 1 for a visualization of the relative robustness. We can prove the following properties for the relative robustness:

Proposition 1. Let x, $y \in F$ be any fixed elements. Let us denote formally $1/t = \infty$ whenever $t = 0$.

(a) The function $w_L(\cdot|y)^{-1} : F \to \mathbb{R} \cup \{\infty\}$ is convex, i.e.,

$$\frac{1}{w_L(tx_1 + (1-t)x_2|y)} \le \frac{t}{w_L(x_1|y)} + \frac{1-t}{w_L(x_2|y)}$$

for any x_1, $x_2 \in F$, $0 \le t \le 1$.

(b) The function $\left(1 - w_L(x|\cdot)\right)^{-1} : L_0 \to \mathbb{R} \cup \{\infty\}$ is concave, i.e.,

$$\frac{1}{1 - w_L(x|sy_1 + (1-s)y_2)} \ge \frac{s}{1 - w_L(x|y_1)} + \frac{1-s}{1 - w_L(x|y_2)}$$

for any y_1, $y_2 \in L_0$, $0 \le s \le 1$.

Proof. Let us prove item (a). Pick x_1, $x_2 \in F$, $y \in F$ and $t \in [0, 1]$ and denote $x = tx_1 + (1-t)x_2$. We may restrict to the case where $w_L(x_r|y) > 0$, $r = 1, 2$. Choose $0 < t_r < w_L(x_r|y)$ such that $z_r := t_r x_r + (1 - t_r)y \in L_0$, $r = 1, 2$. Through simple calculations, one finds that, denoting $t_0 = \left(t/t_1 + (1-t)/t_2\right)^{-1}$ and $s = tt_0/t_1$, one may write

$$L_0 \ni sz_1 + (1-s)z_2 = t_0 x + (1 - t_0)y.$$

This means that $w_L(x|y) \ge t_0$ and, as one lets $t_r \uparrow w_L(x_r|y)$, $r = 1, 2$, the claim is proven.

We go on to proving item (b). Pick y_1, $y_2 \in L_0$ and $s \in [0, 1]$. If $w_L(x|y_1) = 0$ (or $w_L(x|y_2) = 0$), then set $t_1 = 0$ ($t_2 = 0$), otherwise, set $t_1 \in [0, w_L(x|y_1))$ and $t_2 \in [0, w_L(x|y_2))$. Denote $y_0 = sy_1 + (1-s)y_2$ and define $t_0 \in [0, 1]$ through $1 - t_0 = \left(s/(1 - t_1) + (1 - s)/(1 - t_2)\right)^{-1}$. (Note that t_1, $t_2 < 1$.) Let $z_1 \in L_0$, $z_2 \in L_0$ be such that

$$z_1 = t_1 x + (1 - t_1)y_1 \in L_0, \qquad z_2 = t_2 x + (1 - t_2)y_2 \in L_0.$$

Direct calculation shows that we may write $z = t_0 x + (1 - t_0)y$ where

$$z = rz_1 + (1 - r)z_2, \quad r = \frac{s(1 - t_2)}{s(1 - t_2) + (1 - s)(1 - t_1)},$$

and, hence, $z \in L_0$ so that, by definition, $t_0 \le w_L(x|y)$. This amounts to

$$\frac{1}{1 - w_L(x|y)} \ge \frac{s}{1 - t_1} + \frac{1-s}{1 - t_2},$$

and as we let $t_1 \uparrow w_L(x|y_1)$ and $t_2 \uparrow w_L(x|y_2)$, we obtain the desired result. $\square$

Definition 2. For any $x \in F$ let us define

$$w_L(x) = \sup_{y \in L_0} w_L(x|y).$$

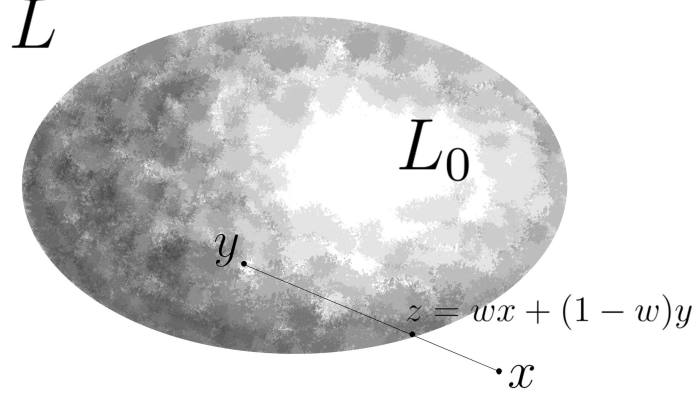We call $w_L(x)$ as the *(absolute) L-robustness* of x.

FIGURE 1. Illustration of the relative robustness function. The points $x \in F$ and $y \in L_0$ are given and typically $x \in L$. The relative robustness is obtained by considering the line segment joining x and y and quantifying the weight w with which x appears in the boundary element z; in this case $w = w_L(x|y)$. One has $z \in L_0$ if L_0 is closed, i.e., the supremum in the evaluation of the absolute robustness is actually obtained.

The absolute robustness $w_L(x)$ measures the overall 'distance' of x to $L_0 = F \setminus L$ in the sense that, whenever L_0 is a compact subset of a locally convex space V, $x \in L_0$ if and only if $w_L(x) = 1$. Moreover, $1 - w_L(x)$ is the greatest amount of random noise from L_0 that x tolerates without being immersed in L_0. If L_0 is absorbing in F, then $w_L(x) > 0$ for all $x \in F$.

Suppose that $x \in F$, y, $z \in L_0$ and $t \in [0,1]$ are such that $z = tx + (1-t)y$. Let us assume that there is an element y' of L_0 on the line connecting x, y and z such that $y' = (1+p)y - px$ for some $p > 0$, i.e., y' is 'behind' y when looked from x. It follows that $z = t'x + (1-t')y'$, where $t' = (t+p)/(1+p) = t + (1-t)p/(1+p) > t$. From this we may conclude that in order to increase $w_L(x|y)$ for a fixed $x \in F$ over $y \in L_0$ we should look for points y of L_0 'furthest away' from x, as is also illustrated in Figure 2. However, the maximizing point (if it exists) is typically not an extreme point of L_0.

We may prove the following:

Proposition 2. The function $\left(w_L(\cdot)\right)^{-1} : F \to \mathbb{R} \cup \{\infty\}$ is convex, i.e., for all x_1, $x_2 \in F$ and $t \in [0,1]$

$$\frac{1}{w_L(tx_1 + (1-t)x_2)} \le \frac{t}{w_L(x_1)} + \frac{1-t}{w_L(x_2)}.$$

Proof. Pick x_1, $x_2 \in F$ and $t \in [0,1]$ and define $x = tx_1 + (1-t)x_2$. Clearly, we may restrict to the case where $w_L(x_1)$, $w_L(x_2) > 0$. Moreover, let $t_1 \in (0, w_L(x_1))$, $t_2 \in (0, w_L(x_2))$ and y_1, $y_2 \in L_0$ be such that

$$z_1 = t_1 x_1 + (1-t_1)y_1 \in L_0, \qquad z_2 = t_2 x_2 + (1-t_2)y_2 \in L_0.$$

Define $t_0 = \left(t/t_1 + (1-t)/t_2\right)^{-1}$, and

$$
\begin{aligned}
y &= \frac{t_0}{1-t_0}\Big(t\frac{1-t_1}{t_1}y_1 + (1-t)\frac{1-t_2}{t_2}y_2\Big), \\
z &= t_0\Big(\frac{t}{t_1}z_1 + \frac{1-t}{t_2}z_2\Big).
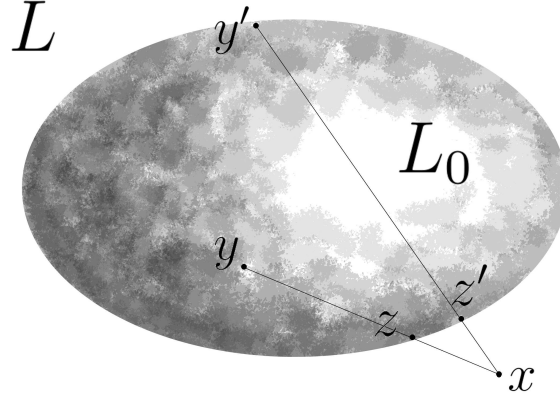\end{aligned}
$$

FIGURE 2. Consider, as is illustrated, points $x \in L$ and y, $y' \in L_0$ situated so that y' is fractionally further away from x than y is in the sense that, when we consider the boundary points z and z' found similarly as in Figure 1, x has a higher weight in z' than in z, i.e., $w(x|y) < w(x|y')$. It follows, as can be seen in the illustration, that these further-away points are located towards the boundary of the 'other end' of L_0 seen from x.

It is easy to check that y is a convex combination of y_1 and y_2 and z is a convex combination of z_1 and z_2, meaning $y \in L_0$ and $z \in L_0$. Furthermore, one may write $t_0 x + (1 - t_0)y = z$, so that $t_0 \leq w_L(x)$ which is equivalent to

$$\frac{1}{w_L(x)} \leq \frac{t}{t_1} + \frac{1-t}{t_2}.$$

Letting $t_1 \uparrow w_L(x_1)$ and $t_2 \uparrow w_L(x_2)$, the claim is proven. $\qquad \square$

In the physical situations, the set of actual selections cannot be described as a convex plane F but rather as a restricted convex subset of a plane, and we are hence interested in distinguishing the useful selections out of the useless selections $y \in L_0$ within restricted sets. Thus, let us fix another convex set K, $L_0 \subset K \subset F$. In most practical cases, the set K will be compact with respect to some locally convex topology of V. Let us restrict the study of the earlier defined relative robustness measures to inputs from K instead of the entire affine subspace F. We may define the following measure for an element of K not being in L_0:

Definition 3. We define the function $w_L^K : K \to \mathbb{R}$,

$$w_L^K(x) = \sup_{y \in K} w_L(x|y), \qquad x \in K,$$

and we call the number $w_L^K(x)$ as the *(absolute) (K, L)-robustness of x*.

The (K, L)-robustness w_L^K measures 'distances' of elements in K from L_0 much in the same way as w_L; especially, in the case where L_0 and K are compact, $x \in K$ is in L_0 if and only if $w_L^K(x) = 1$. Again, $1 - w_L^K(x)$ is the greatest amount of noise from K that x tolerates without being immersed in L_0. For geometric insight into w_L^K and for how w_L^K differs from w_L, see Figure 3. Like $1/w_L$ also $1/w_L^K$ is convex; the proof is essentially the same as the proof for w_L.
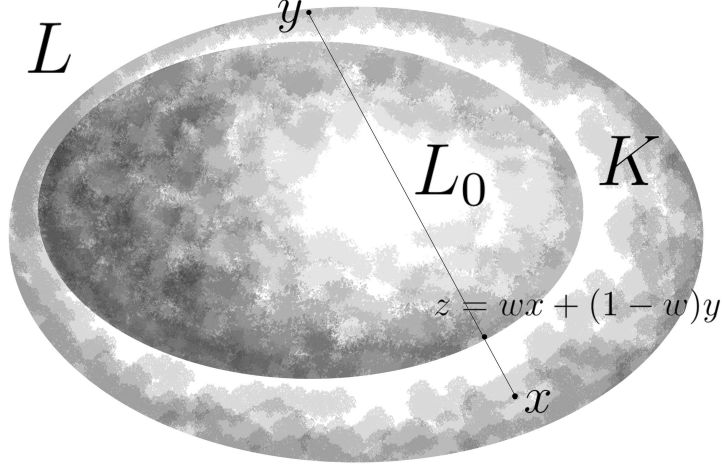
FIGURE 3. The absolute (K, L)-robustness of a point $x \in K \setminus L_0$ (i.e., x is in the outer layer of the highlighted area in the illustration) is obtained by considering the points y of K being fractionally as far away from x as possible in sense clarified in Figure 2. However, now we do not require y to be in L_0. Clearly, the optimizing y is on the boundary of K.

The (K, L)-robustness is typically easier to compute and, hence, we mainly concentrate on this measure in our examples in Section 5.

Remark 1. There is another, largely equivalent way to define the robustness measures. For any $x, y \in F$, let us define $R_L(x|y)$ as the infimum of the numbers $\lambda \in [0, \infty]$ such that $(1 + \lambda)^{-1}(x + \lambda y) \in L_0$, where we set $\inf \emptyset := \infty$. Hence, we can set up the functions $R_L : F \to \mathbb{R} \cup \{\infty\}$ and $R_L^K : K \to \mathbb{R} \cup \{\infty\}$,

$$R_L(x) = \inf_{y \in L_0} R_L(x|y), \qquad R_L^K(x) = \inf_{y \in K} R_L(x|y).$$

It is immediate that $R_L = 1/w_L - 1$ and $R_L^K = 1/w_L^K - 1$. Hence, according to Proposition 2, R_L and, similarly, R_L^K are convex. If L_0 (and, additionally, K) is compact and convex, it follows that $R_L(x) = 0$ (resp. $R_L^K = 0$) if and only if $x \in L_0$. The measures R_L and R_L^K are, thus, more reminiscent of true measures of distance from L_0 than w_L and w_L^K, but because of the ease of calculation and clearer geometric meaning we use the measures w_L and w_L^K, instead. Of course, all our results concerning the measures w_L and w_L^K are transferable into properties of R_L and R_L^K.

The measures R_L and R_L^K have been applied in quantum physics to measure the entanglement of states of combined quantum systems. In this case, K is the set of states of a combined (bipartite) quantum system and L_0 is the subset of separable quantum states, i.e., the closed convex hull of product states $\rho_1 \otimes \rho_2$. The measure R_L, denoted simply R was, in this context, introduced in [24] as the *robustness of entanglement*. In the sequel, we study the measures w_L and w_L^K in the context of quantum incompatibility.

3. THE MATHEMATICAL DESCRIPTION OF BASIC QUANTUM DEVICES

In this section we fix complex and separable Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$; we denote the algebra of bounded operators on $\mathcal{H}$ (respectively $\mathcal{K}$) by $\mathcal{L}(\mathcal{H})$ (respectively $\mathcal{L}(\mathcal{K})$). The unit of $\mathcal{L}(\mathcal{H})$ (the identity operator) is represented by $\mathbb{1}_{\mathcal{H}}$, although we usually omit the subscript if there is no danger of confusion. We denote the set of trace class-operators on $\mathcal{H}$ (respectively on

$\mathcal{K}$) by $\mathcal{T}(\mathcal{H})$ (respectively by $\mathcal{T}(\mathcal{K})$); when endowed with the trace norm the set of trace-class operators becomes a Banach space. We define $\mathcal{S}(\mathcal{H})$ (respectively $\mathcal{S}(\mathcal{K})$) as the set of positive elements of $\mathcal{T}(\mathcal{H})$ (respectively of $\mathcal{T}(\mathcal{K})$) of trace 1.

Furthermore, (Ω, Σ) and (Ω', Σ') will be measurable spaces, i.e., Ω (respectively Ω') is a non-empty set and Σ (respectively Σ') is a σ-algebra of subsets of Ω (respectively Ω'). Additionally we assume that (Ω, Σ) and (Ω', Σ') are standard Borel (i.e., σ-isomorphic to the Borel measurable space of a Borel subset of a Polish space); this makes the notion of post-processing introduced later easier to handle.

The quantum state space of a system described by the Hilbert space $\mathcal{H}$ is identified with $\mathcal{S}(\mathcal{H})$. An (Ω, Σ)-valued quantum observable M is an affine map of $\mathcal{S}(\mathcal{H})$ into the set of probability measures on (Ω, Σ), $\rho \mapsto p_\rho^\mathsf{M}$. The number $p_\rho^\mathsf{M}(X)$ is the probability of registering an outcome from the set $X \in \Sigma$ in a measurement of M when the system is in state ρ.

Hence, an observable is represented by (and, indeed, from now on identified with) a *normalized positive-operator-valued measure* (POVM), i.e., an (Ω, Σ)-valued observable M on a physical system described by $\mathcal{H}$ is a map $\mathsf{M} : \Sigma \to \mathcal{L}(\mathcal{H})$ such that for any state $\rho \in \mathcal{S}(\mathcal{H})$ the function $p_\rho^\mathsf{M} : \Sigma \to \mathbb{R}$, $p_\rho^\mathsf{M}(X) = \mathrm{tr}[\rho \mathsf{M}(X)]$, is a probability measure. This means that, as an operator-valued set function, M is weakly σ-additive, the range $\mathrm{ran}\, \mathsf{M} = \{\mathsf{M}(X) \,|\, X \in \Sigma\}$ consists of positive operators, and $\mathsf{M}(\Omega) = \mathbb{1}_\mathcal{H}$. A particular class of observables, the set of *sharp observables*, is made up of *projection-valued measures* (PVMs) P whose range consists solely of projections. We denote the set of (Ω, Σ)-valued observables of a system described by $\mathcal{H}$ (identified with POVMs) by $\mathbf{Obs}(\Sigma, \mathcal{H})$.

A transformation of a system associated with the Hilbert space $\mathcal{H}$ into a system associated with a possibly different Hilbert space $\mathcal{K}$ is described by an affine map $\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{K})$ that is completely positive. This means that, for the transposed map $\mathcal{E}^* : \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$,

$$\sum_{j,k=1}^n \langle \varphi_j | \mathcal{E}^*(B_j^* B_k) \varphi_k \rangle \geq 0$$

for all $n = 1, 2, \ldots$, $\varphi_1, \ldots, \varphi_n \in \mathcal{H}$, and $B_1, \ldots, B_n \in \mathcal{L}(\mathcal{K})$, and $\mathcal{E}^*(\mathbb{1}_\mathcal{K}) = \mathbb{1}_\mathcal{H}$. Such maps are called as *channels* and we denote the set of channels $\mathcal{E} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{K})$ by $\mathbf{Ch}(\mathcal{H}, \mathcal{K})$.

Let $\mathcal{K}$ be another Hilbert space, $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$, and $\mathsf{N} \in \mathbf{Obs}(\Sigma, \mathcal{K})$. We say that N is *pre-processing of M (by a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$)* if $\mathsf{N} = \mathcal{E}^* \circ \mathsf{M}$. In other words, $p_\rho^\mathsf{N} = p_{\mathcal{E}(\rho)}^\mathsf{M}$. Thus a measurement of N can be implemented by first transforming the system with the channel $\mathcal{E}$ and then measuring M on the transformed system.

Assume now that $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathsf{N} \in \mathbf{Obs}(\Sigma', \mathcal{H})$. Suppose that ρ_0 is a fixed faithful state operator on $\mathcal{H}$ and denote $\mu = p_{\rho_0}^\mathsf{M}$. If there exists a function $\beta : \Sigma' \times \Omega \to \mathbb{R}$ such that

   (i) for any $Y \in \Sigma'$ the function $\beta(Y, \cdot) : \Omega \to \mathbb{R}$ is Σ-measurable and
   (ii) the set function $\beta(\cdot, \omega) : \Sigma' \to \mathbb{R}$ is a probability measure for μ- almost all $\omega \in \Omega$

(i.e., β is a Markov kernel) such that

$$\mathsf{N}(Y) = \int_\Omega \beta(Y, \omega) \mathsf{M}(d\omega)$$

for all $Y \in \Sigma'$, we say that N is a *post-processing of M (with the Markov kernel β)*. We usually write, in this context, $\mathsf{N} = \mathsf{M}^\beta$. Hence, we may measure N by first measuring M and then processing the outcome probability distribution of M by β. For more on post-processing (or coarse-graining) especially in finite-outcome settings, see, [21]. The appropriate generalizations needed in the case involving measurable spaces that are not standard Borel and deeper issues in post-processing are particularly studied in [18, 19].

We say that a map $\Phi : \mathcal{T}(\mathcal{H}) \to \mathcal{T}(\mathcal{K})$ is an *operation* if it is linear, completely positive, and $\mathrm{tr}[\Phi(\rho)] \leq 1$ for all $\rho \in \mathcal{S}(\mathcal{H})$. It follows that an operation is trace-norm continuous. Complete

positivity of Φ means that the (normal) dual map $\Phi^* : \mathcal{L}(\mathcal{K}) \to \mathcal{L}(\mathcal{H})$ is completely positive. We call a weakly σ-additive map $\Sigma \ni X \mapsto \Gamma_X$, where Γ_X is an operation for all $X \in \Sigma$ and $\Gamma_\Omega \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, an *instrument*. Weak σ-additivity means that, for any $B \in \mathcal{L}(\mathcal{K})$ and $T \in \mathcal{T}(\mathcal{H})$, the map $X \mapsto \mathrm{tr}[B\Gamma_X(T)]$ is σ-additive. We denote the set of instruments associated with the measurable space (Ω, Σ) and the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ as above by $\mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$.

An instrument $\Gamma \in \mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$ is a mathematical description of a measurement process; when the system is in the state ρ, $\Gamma_X(\rho)$ is the non-normalized conditional state after the measurement described by Γ conditioned by an outcome being measured in the subset X and $\mathrm{tr}[\Gamma_X(\rho)]$ is the probability of registering an outcome from X when the input state of the system is ρ. Hence, an instrument combines the description of measurement outcome statistics depending on the input state of the system (observable) with the knowledge of the conditional state changes (with the unconditioned state-change Γ_Ω being a channel).

The sets $\mathbf{Obs}(\Sigma, \mathcal{H})$, $\mathbf{Ch}(\mathcal{H}, \mathcal{K})$, and $\mathbf{Ins}(\Sigma, \mathcal{H})$ of observables, channels, and instruments are convex, as they should as measurement device sets of a statistical physical theory. As an example, for Γ, $\Gamma' \in \mathbf{Ins}(\Sigma, \mathcal{H})$ and $t \in [0, 1]$, the convex combination $t\Gamma + (1-t)\Gamma' \in \mathbf{Ins}(\Sigma, \mathcal{H})$ is defined through $\big(t\Gamma + (1-t)\Gamma'\big)_X(\rho) = t\Gamma_X(\rho) + (1-t)\Gamma'_X(\rho)$ for all $X \in \Sigma$ and $\rho \in \mathcal{S}(\mathcal{H})$. It is noteworthy that, unlike in classical physical theories, the convex structures in quantum theory allow typically (uncountably) many decompositions into extreme points for states [6] as well as for measurement devices meaning that a mixture $t\Phi + (1-t)\Psi$ of quantum devices Φ and Ψ with $0 < t < 1$ cannot be considered as an ensemble of devices where the device realized would be, in fact, Φ with probability t and Ψ with probability $1-t$.

4. Quantum compatibility and incompatibility

In this section we give a description of compatibility and the complementary notion of incompatibility of the relevant quantum devices, namely, observables and channels. Incompatibility of quantum observables is a well-known issue; see, e.g., review of the subject in [20] and references therein. The compatibility of other types of devices is dealt with earlier, e.g., in [9, 13], and our discussion here follows the definitions made in these references. Let us fix Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ and the standard Borel value spaces (Ω, Σ) and (Ω', Σ').

We say that observables $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathsf{N} \in \mathbf{Obs}(\Sigma', \mathcal{H})$ are *compatible* or *jointly measurable* if there is a third value space $(\overline{\Omega}, \overline{\Sigma})$ and an observable $\mathsf{G} \in \mathbf{Obs}(\overline{\Sigma}, \mathcal{H})$ such that M and N are post-processings of G. Joint measurability of M and N means that we may determine the outcome statistics of M and N from the outcome statistics of G by classical means (Markov kernels).[1] If a pair of observables is not jointly measurable, we say that the observables are *incompatible*.

Since the value spaces of the observables we study are assumed to be standard Borel, we find an observable $\mathsf{G} \in \mathbf{Obs}(\Sigma \otimes \Sigma', \mathcal{H})$ on the product measurable space $(\Omega \times \Omega', \Sigma \otimes \Sigma')$ for any pair $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathsf{N} \in \mathbf{Obs}(\Sigma', \mathcal{H})$ of jointly measurable observables such that

$$\mathsf{G}(X \times \Omega') = \mathsf{M}(X), \qquad \mathsf{G}(\Omega \times Y) = \mathsf{N}(Y)$$

for all $X \in \Sigma$ and $Y \in \Sigma'$. We call such an observable G as a *joint observable* for M and N. The joint observable of a pair of jointly measurable observables need not be unique. However, if M is an extreme point of $\mathbf{Obs}(\Sigma, \mathcal{H})$ or N is an extreme point of $\mathbf{Obs}(\Sigma', \mathcal{H})$ and M and N are jointly measurable, their joint observable is unique [9].

---

[1]Here is also the connection of joint measurability of observables and non-existence of steering: steering is not present when one party (Alice) cannot steer the other party's (Bob) state outside the range obtained by post-processings of a fixed ensemble of states over a hidden variable with her measurements (observables). Steering will never be realized independently of the initial state if and only if the observables Alice measures are jointly measurable.

Note that any observable $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ is compatible with itself. Indeed, let $(\mathcal{M}, \mathsf{P}, J)$ be a (not necessarily minimal) Naĭmark dilation for M, where $\mathcal{M}$ is a Hilbert space, $\mathsf{P} : \Sigma \to \mathcal{L}(\mathcal{M})$ is a projection valued measure, and $J : \mathcal{H} \to \mathcal{M}$ is an isometry such that $\mathsf{M}(X) = J^*\mathsf{P}(X)J$ for all $X \in \Sigma$. The observable $\mathsf{G} \in \mathbf{Obs}(\Sigma \otimes \Sigma, \mathcal{H})$, $\mathsf{G}(X \times Y) = J^*\mathsf{P}(X \cap Y)J$, $X, Y \in \Sigma$, is clearly a joint observable for the pair (M, M). We say that an observable $\mathsf{T} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ is *trivial*, if there is a probability measure $p : \Sigma \to [0, 1]$ such that $\mathsf{T}(X) = p(X)\mathbb{1}_{\mathcal{H}}$ for all $X \in \Sigma$. It is immediate that the trivial observables are compatible with any other observables.

Whenever $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1 \otimes \mathcal{K}_2)$, we may define its *marginals* $\mathcal{F}_{(j)} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_j)$, $j = 1, 2$, through

$$\mathcal{F}_{(1)} = \mathrm{tr}_{\mathcal{K}_2}[\mathcal{F}(\cdot)], \qquad \mathcal{F}_{(2)} = \mathrm{tr}_{\mathcal{K}_1}[\mathcal{F}(\cdot)]$$

or, equivalently in the dual form,

$$\mathcal{F}^*_{(1)}(A) = \mathcal{F}^*(A \otimes \mathbb{1}_{\mathcal{K}_2}), \qquad \mathcal{F}^*_{(2)}(B) = \mathcal{F}^*(\mathbb{1}_{\mathcal{K}_1} \otimes B)$$

for all $A \in \mathcal{L}(\mathcal{K}_1)$ and $B \in \mathcal{L}(\mathcal{K}_2)$. Clearly, the marginals are channels as well. The marginal channels describe, e.g., the reduced dynamics of subsystems.

We say that $\mathcal{E}_j \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_j)$, $j = 1, 2$, are *compatible* if they are marginals of a channel $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1 \otimes \mathcal{K}_2)$, i.e., $\mathcal{E}_1 = \mathcal{F}_{(1)}$ and $\mathcal{E}_2 = \mathcal{F}_{(2)}$. In this case, we call $\mathcal{F}$ as a *joint channel* of $\mathcal{E}_1$ and $\mathcal{E}_2$. If a pair of channels is not compatible, we say that they are *incompatible*. Compatibility of channels parallels the joint measurability of observables. As with joint measurability, the joint channel of a compatible pair need not be unique. However, a similar sufficient condition for uniqueness can be established as in the case of jointly measurable observables [9]. Due to the deeper non-commutativity of channels, a channel may not be compatible with itself. Indeed, the no-cloning principle can be stated in the form that the identity channel id, $\mathrm{id}(B) = B$, is not compatible with itself. We say that a channel that is not compatible with itself is *self-incompatible*. In a sense, the reason for the fact that (practically) any observable is self-compatible whereas a channel is not is that the output of an observable is classical information that can be copied freely but the quantum output of a channel cannot be copied because of the self-incompatibility of the identity channel.

Finally, we come to the compatibility criterion for a pair of a quantum observable and a channel. Such a pair is compatible if they can be combined in a single measurement, i.e., an instrument. An instrument $\Gamma \in \mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$ has the observable marginal $\Gamma_{(1)} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and the channel marginal $\Gamma_{(2)} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ defined by

$$\Gamma_{(1)}(X) = \Gamma^*_X(\mathbb{1}_{\mathcal{K}}), \qquad \Gamma_{(2)}(\rho) = \Gamma_\Omega(\rho)$$

for all $X \in \Sigma$ and all $\rho \in \mathcal{S}(\mathcal{H})$. The observable $\Gamma_{(1)}$ is the observable whose measurement is realized by the measurement process described by Γ and $\Gamma_{(2)}$ is the unconditioned total state change induced by the measurement.

Compatibility of an observable $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ means that they can be combined in a single measurement, i.e., there is an instrument $\Gamma \in \mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$ such that $\mathsf{M} = \Gamma_{(1)}$ and $\mathcal{E} = \Gamma_{(2)}$, as is highlighted in Figure 4. In this context, we call Γ as a *joint instrument* for M and $\mathcal{E}$. Again, a compatible pair $(\mathsf{M}, \mathcal{E})$ typically has infinitely many joint instruments but one sufficient condition for uniqueness of the joint instrument is the extremality of a marginal. A pair $(\mathsf{M}, \mathcal{E})$ is defined to be *incompatible* if it is not compatible.

4.1. **Robustness of incompatibility.** We fix the sets $\mathbf{Q}_1$ and $\mathbf{Q}_2$ of quantum devices that are either observables or channels, i.e., with fixed Hilbert spaces $\mathcal{H}$, $\mathcal{K}$, and $\mathcal{K}'$ and value spaces (Ω, Σ) and (Ω', Σ'), $\mathbf{Q}_1$ is either $\mathbf{Obs}(\Sigma, \mathcal{H})$ or $\mathbf{Ch}(\mathcal{H}, \mathcal{K})$ and $\mathbf{Q}_2$ is either $\mathbf{Obs}(\Sigma', \mathcal{H})$ or $\mathbf{Ch}(\mathcal{H}, \mathcal{K}')$. We denote the set of compatible pairs within $\mathbf{Q}_1 \times \mathbf{Q}_2$ by $\mathbf{Comp}$, i.e., $\Phi_1 \in \mathbf{Q}_1$
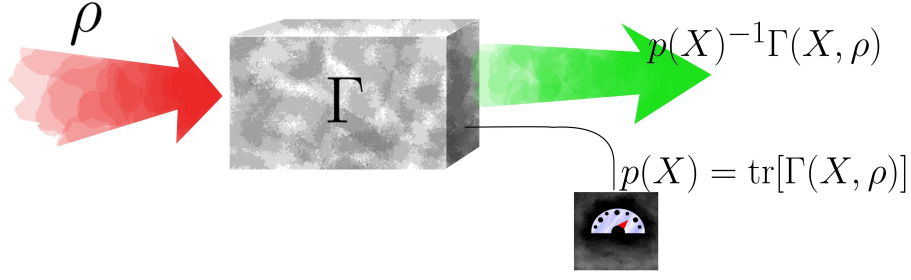
FIGURE 4. Illustration of an instrument. The state entering the measurement device represented by the instrument Γ is ρ. The instrument has the statistics arm (the lower branch right of the instrument box in the illustration) and the state change arm (the upper branch right of the instrument box). When a value is detected with certainty in the set X in the statistics arm, which happens with probability $p(X)$, the state change arm gives the conditional state $\rho_X = p(X)^{-1}\Gamma_X(\rho)$. When the state changes are neglected, the statistics arm reduces to the observable $\Gamma_{(1)}$, and when the statistics are ignored, the state-change arm reduces to the channel $\Gamma_{(2)}$. All joint maps of compatible quantum devices can be visualized in the same way as input-output processors with multiple outputs. Each of the original compatible devices is obtained when all the other output arms are ignored except for the one associated with the particular device.

and $\Phi_2 \in \mathbf{Q}_2$ are compatible if and only if $(\Phi_1, \Phi_2) \in \mathbf{Comp}$. The set $\mathbf{Q}_1 \times \mathbf{Q}_2$ is endowed with a natural convex structure by defining for any (Φ_1, Φ_2), $(\Psi_1, \Psi_2) \in \mathbf{Q}_1 \times \mathbf{Q}_2$ and $t \in [0, 1]$ the combination $t(\Phi_1, \Phi_2) + (1 - t)(\Psi_1, \Psi_2) = (t\Phi_1 + (1 - t)\Psi_1, t\Phi_2 + (1 - t)\Psi_2)$. Whatever the sets of devices involved, the reader may easily check that $\mathbf{Comp}$ is a convex subset of $\mathbf{Q}_1 \times \mathbf{Q}_2$.

Denote by L the relative complement of $\mathbf{Comp}$ with respect to the minimal affine subspace containing $\mathbf{Comp}$, which coincides with the minimal affine subspace containing $\mathbf{Q}_1 \times \mathbf{Q}_2$; see Remark 2 for a proof of this fact. The product $\mathbf{Q}_1 \times \mathbf{Q}_2$ we denote by K. We may define the robustness measures $w_L(\cdot|\cdot)$, w_L, and w_L^K introduced in general form in Section 2 for the set of compatible pairs $\mathbf{Comp}$. For simplicity, we denote $w_L^K(\cdot|\cdot) =: w(\cdot|\cdot)$, $w_L =: w$, and $w_L^K =: W$. Moreover, for any Φ_1, $\Psi_1 \in \mathbf{Q}_1$ and Φ_2, $\Psi_2 \in \mathbf{Q}_2$, we simplify our notations:

$$
\begin{aligned}
w\big((\Phi_1, \Phi_2)|(\Psi_1, \Psi_2)\big) &=: w(\Phi_1, \Phi_2|\Psi_1, \Psi_2), \\
w\big((\Phi_1, \Phi_2)\big) &=: w(\Phi_1, \Phi_2), \\
W\big((\Phi_1, \Phi_2)\big) &=: W(\Phi_1, \Phi_2).
\end{aligned}
$$

When $\mathbf{Q}_1 = \mathbf{Q}_2 = \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, we denote $W(\mathcal{E}, \mathcal{E}) = W(\mathcal{E})$; this quantity we call as the *robustness of self-incompatibility of* $\mathcal{E}$.

In the case where $\mathbf{Q}_1 = \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathbf{Q}_2 = \mathbf{Ch}(\mathcal{H}, \mathcal{H})$, the quantity $\mathbf{Obs}(\Sigma, \mathcal{H}) \ni \mathsf{M} \mapsto W(\mathsf{M}, \mathrm{id})$, where id stands for the identity channel in $\mathbf{Ch}(\mathcal{H}, \mathcal{H})$, measures how well an approximate version of M can be measured while disturbing the system as little as possible. Equivalently, one may think of the number $W(\mathsf{M}, \mathrm{id})$ as the measure of how disturbing any measurement of the observable M inherently is. In Section 5.3, this quantity is calculated in the case of finite-dimensional $\mathcal{H}$ for any rank-1 sharp observable.

Remark 2. In [4], it was shown that, whenever $t \leq 1/2$, $(\mathsf{A}, \mathsf{B}) \in \mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Obs}(\Sigma', \mathcal{H})$, and $(\mathsf{S}, \mathsf{T}) \in \mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Obs}(\Sigma', \mathcal{H})$ is a pair of trivial observables, then $t(\mathsf{A}, \mathsf{B}) + (1 - t)(\mathsf{S}, \mathsf{T}) \in$

$\mathbf{JM}(\Sigma, \Sigma', \mathcal{H})$, so that

$$W(\mathsf{A}, \mathsf{B}) \geq w(\mathsf{A}, \mathsf{B}|\mathsf{S}, \mathsf{T}) \geq \frac{1}{2}.$$

Hence, $1/2$ is a global lower bound for the robustness W. Indeed, the same reasoning as that in [4] also applies to the compatibility questions involving other devices than observables as well as to the case of continuous observables not discussed in [4]. Let us study, e.g., the case of observable-channel pairs, i.e., $\mathbf{Q}_1 = \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathbf{Q}_2 = \mathbf{Ch}(\mathcal{H}, \mathcal{K})$. Pick any probability measure p on (Ω, Σ) and any state $\sigma \in \mathcal{S}(\mathcal{K})$ defining the associated trivial observable $\mathsf{T}_p = p(\cdot)\mathbb{1}_{\mathcal{H}}$ and the constant channel $\mathcal{T}_\sigma : \rho \mapsto \sigma$. With any $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$, $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, and $t \in [0, 1]$, one may set up the instrument $\Gamma \in \mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$,

$$\Gamma_X(\rho) = t\,\mathrm{tr}[\rho\mathsf{M}(X)]\sigma + (1-t)p(X)\mathcal{E}(\rho), \qquad \rho \in \mathcal{S}(\mathcal{H}).$$

The observable marginal of Γ is easily found to be $t\mathsf{M} + (1-t)\mathsf{T}_p$ and the channel marginal is $t\mathcal{T}_\sigma + (1-t)\mathcal{E}$. Especially it follows that fixing any pair $(\mathsf{T}_p, \mathcal{T}_\sigma)$ like that above, one has that $\frac{1}{2}(\mathsf{M} + \mathsf{T}_p)$ and $\frac{1}{2}(\mathcal{E} + \mathcal{T}_\sigma)$ are compatible. Similar result is easily proven for channel-channel pairs. It thus follows that there is $x_0 \in \mathbf{Comp}$ such that $\frac{1}{2}(x + x_0) \in \mathbf{Comp}$ for all $x \in \mathbf{Q}_1 \times \mathbf{Q}_2$ implying $W \geq 1/2$ for any device pair

It also follows that the minimal affine subspace F' containing $\mathbf{Comp}$ coincides with the minimal affine subspace F containing $\mathbf{Q}_1 \times \mathbf{Q}_2$. Indeed, trivially, $F' \subset F$. For the reversed inclusion, let us pick any $x_0 \subset \mathbf{Comp}$ such that $\frac{1}{2}(x + x_0) \in \mathbf{Comp}$ for all $x \in \mathbf{Q}_1 \times \mathbf{Q}_2$. Let $z \in F$ meaning that there are $x, y \in \mathbf{Q}_1 \times \mathbf{Q}_2$ and $\lambda \geq 0$ such that $z = x_0 + \lambda(x - y)$. Defining $x' := \frac{1}{2}(x + x_0) \in \mathbf{Comp}$ and $y' := \frac{1}{2}(y + x_0) \in \mathbf{Comp}$, it follows $z = x_0 + 2\lambda(x' - y') \in F'$ implying $F \subset F'$.

Following [14], in the case for observables operating in a d-dimensional ($d < \infty$) Hilbert space one can give an even tighter bound for robustness of incompatibility for observable pairs:

$$W(\mathsf{A}, \mathsf{B}) \geq \frac{2 + d}{2(1 + d)}.$$

One easily sees, using similar techniques as in [14] that the above inequality holds also for the robustness measures W involving quantum devices other than observables.

4.2. **Ordering properties.** In this subsection we discuss some of the special features of the robustness measures for incompatibility. We will find out that the robustness measure behaves monotonically under certain orderings of device pairs. In the sequel, whenever (Φ, Ψ) is a device pair and we write $W(\Phi, \Psi)$ (or $w(\Phi, \Psi)$) we implicitly assume that the base set $\mathbf{Q}_1 \times \mathbf{Q}_2 \ni (\Phi, \Psi)$ contains compatible pairs, i.e., if Φ and Ψ are observables, they operate in the same Hilbert space, if Φ and Ψ are channels, their input spaces coincides, and if Φ is an observable and Ψ is a channel, then Φ operates in the input space of the channel Ψ.

Let us fix the Hilbert spaces $\mathcal{H}$ and $\mathcal{K}$ and four standard Borel measurable spaces (Ω, Σ), (Ω', Σ'), $(\tilde{\Omega}, \tilde{\Sigma})$, and $(\tilde{\Omega}', \tilde{\Sigma}')$. We denote the subset of those pairs $(\mathsf{M}, \mathsf{N}) \in \mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Obs}(\Sigma', \mathcal{H})$ that are jointly measurable by $\mathbf{JM}(\Sigma, \Sigma', \mathcal{H})$. We define the corresponding sets of compatible observable pairs also for other pairs of value spaces.

Suppose that $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{K})$ is a pre-processing of $\mathsf{A} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathsf{N} \in \mathbf{Obs}(\Sigma', \mathcal{K})$ is a pre-processing of $\mathsf{B} \in \mathbf{Obs}(\Sigma', \mathcal{H})$, both with the *same* channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, i.e., $\mathsf{M} = \mathcal{E}^* \circ \mathsf{A}$ and $\mathsf{N} = \mathcal{E}^* \circ \mathsf{B}$. In this case, we write $(\mathsf{M}, \mathsf{N}) \leq_{\mathrm{prae}} (\mathsf{A}, \mathsf{B})$. The relation $\leq_{\mathrm{prae}}$ is a pre-order in the set of observable pairs. We denote $(\mathsf{A}, \mathsf{B}) =_{\mathrm{prae}} (\mathsf{M}, \mathsf{N})$ if $(\mathsf{M}, \mathsf{N}) \leq_{\mathrm{prae}} (\mathsf{A}, \mathsf{B})$ and $(\mathsf{A}, \mathsf{B}) \leq_{\mathrm{prae}} (\mathsf{M}, \mathsf{N})$ and say that the pairs (A, B) and (M, N) are *pre-processing equivalent.*

Another pre-order in the set of observable pairs is defined by post-processing: If $\tilde{\mathsf{M}} \in \mathbf{Obs}(\tilde{\Sigma}, \mathcal{H})$ is a post-processing of $\mathsf{A} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\tilde{\mathsf{N}} \in \mathbf{Obs}(\tilde{\Sigma}', \mathcal{H})$ is a post-processing of $\mathsf{B} \in \mathbf{Obs}(\Sigma', \mathcal{H})$ with possibly different Markov kernels, we write $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \leq_{\mathrm{post}} (\mathsf{A}, \mathsf{B})$. We

define the *post-processing equivalence* $=_{\text{post}}$ for pairs of observables in the same way as the pre-processing equivalence.

Suppose that $(\mathsf{A}, \mathsf{B}) \in \mathbf{JM}(\Sigma, \Sigma', \mathcal{H})$ have a joint observable G. Then $\mathcal{E}^* \circ \mathsf{G}$ is a joint observable for $(\mathcal{E}^* \circ \mathsf{A}, \mathcal{E}^* \circ \mathsf{B})$ for any channel $\mathcal{E}$. This means that, if the pair (A, B) is jointly measurable and $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \leq_{\text{prae}} (\mathsf{A}, \mathsf{B})$, then $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}})$ is jointly measurable. Moreover, it follows immediately from the definition of joint measurability that, if (A, B) is jointly measurable and $(\mathsf{M}, \mathsf{N}) \leq_{\text{post}} (\mathsf{A}, \mathsf{B})$, then (M, N) is jointly measurable.

We may show that the W-robustness measure of incompatibility has the following properties. With a slight modification of the proof given here, one easily shows that these results also hold for w.

Theorem 1. Let (A, B), (M, N), and $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}})$ be observable pairs.

   (a) If $(\mathsf{M}, \mathsf{N}) \leq_{\text{prae}} (\mathsf{A}, \mathsf{B})$, then $W(\mathsf{M}, \mathsf{N}) \geq W(\mathsf{A}, \mathsf{B})$, and if $(\mathsf{M}, \mathsf{N}) =_{\text{prae}} (\mathsf{A}, \mathsf{B})$, then $W(\mathsf{M}, \mathsf{N}) = W(\mathsf{A}, \mathsf{B})$.

   (b) If $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \leq_{\text{post}} (\mathsf{A}, \mathsf{B})$, then $W(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \geq W(\mathsf{A}, \mathsf{B})$, and if $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) =_{\text{post}} (\mathsf{A}, \mathsf{B})$, then $W(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) = W(\mathsf{A}, \mathsf{B})$.

Proof. Clearly, if $W(\mathsf{A}, \mathsf{B}) = 0$, the first claim in item (a) needs no proof. Assume, hence, that $W(\mathsf{A}, \mathsf{B}) > 0$; in fact, according to Remark 2, the robustness measure is always bounded from below by $1/2$. Suppose that $(\mathsf{M}, \mathsf{N}) \leq_{\text{prae}} (\mathsf{A}, \mathsf{B})$, the pre-processing carried out by a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ and $t < W(\mathsf{A}, \mathsf{B})$, and let $(\mathsf{A}_1, \mathsf{B}_1) \in \mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Obs}(\Sigma', \mathcal{H})$ and $(\mathsf{A}_2, \mathsf{B}_2) \in \mathbf{JM}(\Sigma, \Sigma', \mathcal{H})$ be such that

$$t(\mathsf{A}, \mathsf{B}) + (1 - t)(\mathsf{A}_1, \mathsf{B}_1) = (\mathsf{A}_2, \mathsf{B}_2).$$

Let now $\mathsf{M}_r = \mathcal{E}^* \circ \mathsf{A}_r$ and $\mathsf{N}_r = \mathcal{E}^* \circ \mathsf{B}_r$ for $r = 1, 2$. Hence, $(\mathsf{M}_2, \mathsf{N}_2)$ is a compatible pair. It follows immediately (as one may check) that one may write

$$t(\mathsf{M}, \mathsf{N}) + (1 - t)(\mathsf{M}_1, \mathsf{N}_1) = (\mathsf{M}_2, \mathsf{N}_2),$$

from which the first claim of item (a) follows as one lets $t \uparrow W(\mathsf{A}, \mathsf{B})$. The second claim of item (a) follows from symmetry.

In the proof of item (b), we may again restrict to the case where $W(\mathsf{A}, \mathsf{B}) > 0$. Assume now that $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \leq_{\text{post}} (\mathsf{A}, \mathsf{B})$, so that there are Markov kernels β and γ such that $\tilde{\mathsf{M}} = \mathsf{A}^\beta$ and $\tilde{\mathsf{N}} = \mathsf{B}^\gamma$, and $t \leq W(\mathsf{A}, \mathsf{B})$, and let $(\mathsf{A}_1, \mathsf{B}_1) \in \mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Obs}(\Sigma', \mathcal{H})$ and $(\mathsf{A}_2, \mathsf{B}_2) \in \mathbf{JM}(\Sigma, \Sigma', \mathcal{H})$ be as above. Now, $(\mathsf{A}_2^\beta, \mathsf{B}_2^\gamma)$ is a compatible pair, and one may write

$$t(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) + (1 - t)(\mathsf{A}_1^\beta, \mathsf{B}_1^\gamma) = (\mathsf{A}_2^\beta, \mathsf{B}_2^\gamma)$$

proving the first claim of item (b) as one lets $t \uparrow W(\mathsf{A}, \mathsf{B})$. The second claim of item (b) is proven by symmetry again. $\qquad\square$

Especially, if the observables $\mathsf{A} \in \mathbf{Obs}(\Sigma, \mathcal{H})$ and $\mathsf{B} \in \mathbf{Obs}(\Sigma', \mathcal{H})$ are post-processing maximal (i.e., rank-1 observables), then the pair (A, B) minimizes the robustness measure W, i.e., a pair of post-processing maximal observables require the greatest amount of noise to be added in order to be rendered jointly measurable.

Let now $\mathcal{H}$, $\mathcal{H}'$, $\mathcal{K}_1$, $\mathcal{K}_2$, $\mathcal{K}_1'$, and $\mathcal{K}_2'$ be Hilbert spaces. We denote, e.g., for the spaces $\mathcal{H}$, $\mathcal{K}_1$, and $\mathcal{K}_2$, the set of compatible pairs in $\mathbf{Ch}(\mathcal{H}, \mathcal{K}_1) \times \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$ by $\mathbf{Comp}(\mathcal{H}, \mathcal{K}_1, \mathcal{K}_2)$.

For channels $\mathcal{E}' \in \mathbf{Ch}(\mathcal{H}', \mathcal{K}_1)$, $\mathcal{F}' \in \mathbf{Ch}(\mathcal{H}', \mathcal{K}_2)$, $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1)$, and $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$, we denote $(\mathcal{E}', \mathcal{F}') \leq_{\text{prae}} (\mathcal{E}, \mathcal{F})$ if there is a channel $\mathcal{G} \in \mathbf{Ch}(\mathcal{H}', \mathcal{H})$ such that $\mathcal{E}' = \mathcal{E} \circ \mathcal{G}$ and $\mathcal{F}' = \mathcal{F} \circ \mathcal{G}$. This gives rise to the partial order $\leq_{\text{prae}}$ associated to pre-processing of channel pairs. We denote by $=_{\text{prae}}$ the corresponding equivalence relation. Suppose that the $(\mathcal{E}, \mathcal{F}) \in \mathbf{Comp}(\mathcal{H}, \mathcal{K}_1, \mathcal{K}_2)$ has the joint channel $\mathcal{M} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1 \otimes \mathcal{K}_2)$ and pick $\mathcal{G} \in \mathbf{Ch}(\mathcal{H}', \mathcal{H})$. It follows that also the pair $(\mathcal{E} \circ \mathcal{G}, \mathcal{F} \circ \mathcal{G})$ is compatible since it has (among others) the joint channel

$\mathcal{M} \circ \mathcal{G}$. This means that, whenever the pair $(\mathcal{E}, \mathcal{F})$ is compatible and $(\mathcal{E}', \mathcal{F}') \leq_{\mathrm{prae}} (\mathcal{E}, \mathcal{F})$, then also $(\mathcal{E}', \mathcal{F}')$ is compatible.

When $\mathcal{E}' \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1')$, $\mathcal{F}' \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2')$, $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1)$, and $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$, we denote $(\mathcal{E}', \mathcal{F}') \leq_{\mathrm{post}} (\mathcal{E}, \mathcal{F})$ if there are channels $\mathcal{A} \in \mathbf{Ch}(\mathcal{K}_1, \mathcal{K}_1')$ and $\mathcal{B} \in \mathbf{Ch}(\mathcal{K}_2, \mathcal{K}_2')$ such that $\mathcal{E}' = \mathcal{A} \circ \mathcal{E}$ and $\mathcal{F}' = \mathcal{B} \circ \mathcal{F}$. We denote the equivalence relation corresponding to the partial order $\leq_{\mathrm{post}}$ by $=_{\mathrm{post}}$. If $(\mathcal{E}, \mathcal{F}) \in \mathbf{Comp}(\mathcal{H}, \mathcal{K}_1, \mathcal{K}_2)$ has the joint channel $\mathcal{M} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1 \otimes \mathcal{K}_2)$ and we choose $\mathcal{A} \in \mathbf{Ch}(\mathcal{K}_1, \mathcal{K}_1')$ and $\mathcal{B} \in \mathbf{Ch}(\mathcal{K}_2, \mathcal{K}_2')$ we may define the joint channel $(\mathcal{A} \otimes \mathcal{B}) \circ \mathcal{M}$ for the pair $(\mathcal{A} \circ \mathcal{E}, \mathcal{B} \circ \mathcal{F})$. Thus, whenever the pair $(\mathcal{E}, \mathcal{F})$ is compatible and $(\mathcal{E}', \mathcal{F}') \leq_{\mathrm{post}} (\mathcal{E}, \mathcal{F})$, then also $(\mathcal{E}', \mathcal{F}')$ is compatible.

As for observables, we may easily prove the following (the robustness measure w possesses the same properties):

Theorem 2. Let $(\mathcal{E}, \mathcal{F})$, $(\mathcal{C}, \mathcal{D})$, and $(\mathcal{C}', \mathcal{D}')$ be channel pairs.

(a) If $(\mathcal{C}, \mathcal{D}) \leq_{\mathrm{prae}} (\mathcal{E}, \mathcal{F})$, then $W(\mathcal{C}, \mathcal{D}) \geq W(\mathcal{E}, \mathcal{F})$, and if $(\mathcal{C}, \mathcal{D}) =_{\mathrm{prae}} (\mathcal{E}, \mathcal{F})$, then $W(\mathcal{C}, \mathcal{D}) = W(\mathcal{E}, \mathcal{F})$.

(b) If $(\mathcal{C}', \mathcal{D}') \leq_{\mathrm{post}} (\mathcal{E}, \mathcal{F})$, then $W(\mathcal{C}', \mathcal{D}') \geq W(\mathcal{E}, \mathcal{F})$, and if $(\mathcal{C}', \mathcal{D}') =_{\mathrm{post}} (\mathcal{E}, \mathcal{F})$, then $W(\mathcal{C}', \mathcal{D}') = W(\mathcal{E}, \mathcal{F})$.

Thus, especially, we have $W(\mathcal{E}, \mathcal{F}) \geq W(\mathrm{id})$ for any $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1)$ and $\mathcal{F} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$, where $\mathrm{id} \in \mathbf{Ch}(\mathcal{H}, \mathcal{H})$ is the identity channel, i.e., the pair $(\mathrm{id}, \mathrm{id})$ is the most incompatible pair of channels with respect to the robustness measures. The robustness measure W (as well as w) attains the same minimal value at any channel pair in the post-processing equivalence class determined by the identity channel pair $(\mathrm{id}, \mathrm{id})$, $\mathrm{id} \in \mathbf{Ch}(\mathcal{H}, \mathcal{H})$. Clearly, a pair $(\mathcal{E}, \mathcal{F}) \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1) \times \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$ is in this equivalence class when they are left-invertible by channels, i.e., there are channels $\mathcal{A} \in \mathbf{Ch}(\mathcal{K}_1, \mathcal{H})$ and $\mathcal{B} \in \mathbf{Ch}(\mathcal{K}_2, \mathcal{H})$ such that $\mathcal{A} \circ \mathcal{E} = \mathcal{B} \circ \mathcal{F} = \mathrm{id}$. From now on, we call such channels *decodable*. As a special case of [17, Corollary 1], when $\mathcal{H}$ and $\mathcal{K}$ are finite dimensional, a channel $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ is decodable if and only if there is a Hilbert space $\mathcal{K}_0$, a unitary operator $U : \mathcal{H} \otimes \mathcal{K}_0 \to \mathcal{K}$, and a positive trace-1 operator T on $\mathcal{K}_0$ such that $\mathcal{E}(\rho) = U(\rho \otimes T)U^*$ for all $\rho \in \mathcal{S}(\mathcal{H})$. It follows that a channel with unitarily equivalent input and output spaces is decodable if and only if it is a unitary channel, i.e., of the form $\rho \mapsto U\rho U^*$ with a unitary operator U. The decodable channels posses essentially the same properties with respect to the robustness measures as the identity channel.

Let us fix Hilbert spaces $\mathcal{H}$, $\mathcal{H}'$, $\mathcal{K}$, and $\mathcal{K}'$ and standard Borel spaces (Ω, Σ) and (Ω', Σ'). We denote, e.g., for $\mathcal{H}$, $\mathcal{K}$, and Σ, by $\mathbf{Comp}(\Sigma, \mathcal{H}, \mathcal{K})$ the set of compatible pairs in $\mathbf{Obs}(\Sigma, \mathcal{H}) \times \mathbf{Ch}(\mathcal{H}, \mathcal{K})$.

We denote $(\mathsf{M}', \mathcal{E}') \leq_{\mathrm{prae}} (\mathsf{M}, \mathcal{E})$ for $\mathsf{M}' \in \mathbf{Obs}(\Sigma, \mathcal{H}')$, $\mathcal{E}' \in \mathbf{Ch}(\mathcal{H}', \mathcal{K})$, $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$, and $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ if there is $\mathcal{G} \in \mathbf{Ch}(\mathcal{H}', \mathcal{H})$ such that $\mathsf{M}' = \mathcal{G}^* \circ \mathsf{M}$ and $\mathcal{E}' = \mathcal{E} \circ \mathcal{G}$. Again, the equivalence relation corresponding to the partial order $\leq_{\mathrm{prae}}$ is $=_{\mathrm{prae}}$. It is easy to see that, if $(\mathsf{M}, \mathcal{E})$ is a compatible pair and $(\mathsf{M}', \mathcal{E}') \leq_{\mathrm{prae}} (\mathsf{M}, \mathcal{E})$, then $(\mathsf{M}', \mathcal{E}')$ is compatible as well.

If $\mathsf{M}' = \mathsf{M}^\beta$ and $\mathcal{E}' = \mathcal{B} \circ \mathcal{E}$, where $\mathsf{M}' \in \mathbf{Obs}(\Sigma', \mathcal{H})$, $\mathcal{E}' \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}')$, $\mathsf{M} \in \mathbf{Obs}(\Sigma, \mathcal{H})$, and $\mathcal{E} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, for a channel $\mathcal{B} \in \mathbf{Ch}(\mathcal{K}, \mathcal{K}')$ and a Markov kernel $\beta : \Sigma' \times \Omega \to \mathbb{R}$, we denote $(\mathsf{M}', \mathcal{E}') \leq_{\mathrm{post}} (\mathsf{M}, \mathcal{E})$. The equivalence relation associated with the partial order $\leq_{\mathrm{post}}$ is denoted by $=_{\mathrm{post}}$. Suppose that $(\mathsf{M}, \mathcal{E}) \in \mathbf{Comp}(\Sigma, \mathcal{H}, \mathcal{K})$ has the joint instrument $\Gamma \in \mathbf{Ins}(\Sigma, \mathcal{H}, \mathcal{K})$ and pick $\mathcal{B} \in \mathbf{Ch}(\mathcal{K}, \mathcal{K}')$ and a Markov kernel $\beta : \Sigma' \times \Omega \to \mathbb{R}$. It is straight-forward to check that $\Gamma' \in \mathbf{Ins}(\Sigma', \mathcal{H}, \mathcal{K}')$,

$$\Gamma'_Y(\rho) = \mathcal{B}\left(\int_\Omega \beta(Y, \omega) \Gamma_{d\omega}(\rho) \right), \qquad Y \in \Sigma', \quad \rho \in \mathcal{S}(\mathcal{H}),$$

is a joint instrument for $(\mathsf{M}^\beta, \mathcal{B} \circ \mathcal{E})$ implying that, whenever the pair $(\mathsf{M}, \mathcal{E})$ is compatible and $(\mathsf{M}', \mathcal{E}') \leq_{\mathrm{post}} (\mathsf{M}, \mathcal{E})$, then $(\mathsf{M}', \mathcal{E}')$ is compatible as well.

Again, one easily proves the following properties (which also hold for w):

Theorem 3. Let $(\mathsf{M}, \mathcal{E})$, $(\mathsf{N}, \mathcal{F})$, and $(\mathsf{N}', \mathcal{F}')$ be observable-channel pairs.

  (a) If $(\mathsf{N}, \mathcal{F}) \leq_{\mathrm{prae}} (\mathsf{M}, \mathcal{E})$, then $W(\mathsf{N}, \mathcal{F}) \geq W(\mathsf{M}, \mathcal{E})$, and if $(\mathsf{N}, \mathcal{F}) =_{\mathrm{prae}} (\mathsf{M}, \mathcal{E})$, then $W(\mathsf{N}, \mathcal{F}) = W(\mathsf{M}, \mathcal{E})$.
  (b) If $(\mathsf{N}', \mathcal{F}') \leq_{\mathrm{post}} (\mathsf{M}, \mathcal{E})$, then $W(\mathsf{N}', \mathcal{F}') \geq W(\mathsf{M}, \mathcal{E})$, and if $(\mathsf{N}', \mathcal{F}') =_{\mathrm{post}} (\mathsf{M}, \mathcal{E})$, then $W(\mathsf{N}', \mathcal{F}') = W(\mathsf{M}, \mathcal{E})$.

Theorems 1, 2, and 3 together tell that instead of considering robustness measures as functions on individual device pairs, they can be defined on pre- or post-processing equivalence classes. The partial orders invoked by pre- and post-processing in the set of observables and their meaning are studied, e.g., in [3, 11]. The results of this section also tell that the measure $R := 1/W - 1$ (as well as $1/w - 1$) is an incompatibility monotone from the perspective of [12]. The operations defining the preorders $\leq_{\mathrm{prae}}$ and $\leq_{\mathrm{post}}$, common pre-processing and independent bipartite post-processing, can be naturally viewed as compatibility non-decreasing maps and any incompatibility measure should naturally behave monotonously under these operations. Monotonicity under $\leq_{\mathrm{prae}}$ has been required already in [12, 22].

5. Examples

In the remainder of this article, we calculate the robustness of incompatibility W for three special cases: the finite dimensional Weyl pair, the pair of decodable, hence especially unitary, channels, and the pair consisting of a rank-1 sharp observable (von Neumann observable) and a decodable channel. In each case, the quantity $R = 1/W - 1$ measures how well the pair resists joining under noise. Hence, in the first case, we essentially determine the overall resistance to joint measuring of a finite-dimensional 'position-momentum' pair. In the second case, we find how well (or how poorly) we may approximately combine a pair of decodable channels in a single channel. The third case enlightens the issue of how close can the total state change associated with an approximate measurement of a von Neumann observable be to an information-preserving channel

5.1. **Robustness of incompatibility for a sharp Weyl pair.** In this section, we calculate the robustness of incompatibility for a particular pair of incompatible observables: a finite-dimensional Weyl pair. Let us fix a d-dimensional Hilbert space $\mathcal{H}$ ($d < \infty$) which has the orthonormal base $\{\varphi_j\}_{j \in \mathbb{Z}_d}$. We denote $\langle j, k \rangle = e^{i2\pi jk/d}$ for all j, $k \in \mathbb{Z}_d$ and define the linear operator $\mathcal{F} \in \mathcal{L}(\mathcal{H})$ through

$$(5.1) \qquad \mathcal{F}\varphi_j = \frac{1}{\sqrt{d}} \sum_{i \in \mathbb{Z}_d} \overline{\langle i, j \rangle} \varphi_i, \qquad j \in \mathbb{Z}_d.$$

This operator is the Fourier-operator and its adjoint is defined through

$$\mathcal{F}^*\varphi_j = \frac{1}{\sqrt{d}} \sum_{i \in \mathbb{Z}_d} \langle i, j \rangle \varphi_i, \qquad j \in \mathbb{Z}_d.$$

For simplicity, we denote by $\mathbf{Obs}_d$ the set of observables operating in $\mathcal{H}$ whose value space is $\mathbb{Z}_d$ (equipped with its power set as the σ-algebra). Hence, an observable $\mathsf{M} \in \mathbf{Obs}_d$ is defined by the values $\mathsf{M}(\{j\}) := \mathsf{M}_j \in \mathcal{L}(\mathcal{H})$, $j \in \mathbb{Z}_d$, and we write $\mathsf{M} = (\mathsf{M}_j)_{j \in \mathbb{Z}_d}$. We denote the set of compatible pairs in $\mathbf{Obs}_d \times \mathbf{Obs}_d$ by $\mathbf{JM}_d$. We denote the set of $\mathbb{Z}_d \times \mathbb{Z}_d$-valued observables operating in $\mathcal{H}$ (the possible joint observables for the compatible pairs $(\mathsf{M}, \mathsf{N}) \in \mathbf{JM}_d$) by $\mathbf{Obs}_{d \times d}$. When $\mathsf{G} \in \mathbf{Obs}_{d \times d}$, we set $\mathsf{G}_{j,k} := \mathsf{G}(\{(j, k)\})$ for all j, $k \in \mathbb{Z}_d$.

We fix another orthonormal basis $\{\psi_k\}_{k \in \mathbb{Z}_d}$ by setting $\psi_k = \mathcal{F}^* \varphi_k$. It follows that $\langle \varphi_j | \psi_k \rangle = d^{-1/2}$ for all j, $k \in \mathbb{Z}_d$, so that the bases $\{\varphi_j\}$ and $\{\psi_k\}$ are an example of a pair of mutually unbiased bases. Let us denote

$$\mathsf{Q}_j := |\varphi_j\rangle\langle\varphi_j|, \quad \mathsf{P}_k := |\psi_k\rangle\langle\psi_k|, \qquad j, k \in \mathbb{Z}_d$$

and define the sharp observables $\mathsf{Q} := (\mathsf{Q}_j)_{j \in \mathbb{Z}_d} \in \mathbf{Obs}_d$ and $\mathsf{P} := (\mathsf{P}_k)_{k \in \mathbb{Z}_d} \in \mathbf{Obs}_d$.

For each q, $p \in \mathbb{Z}_d$, we may define the operators U_q, V_p, $W_{q,p} \in \mathcal{L}(\mathcal{H})$ through

$$(5.2) \qquad\qquad\qquad U_q \varphi_j = \varphi_{j+q},$$

$$(5.3) \qquad\qquad\qquad V_p \varphi_j = \langle j, p\rangle \varphi_j,$$

$$(5.4) \qquad\qquad\qquad W_{q,p} = U_q V_p,$$

where the sums and differences are considered as cyclic on $\mathbb{Z}_d$. Thus $(q, p) \mapsto W_{q,p}$ is a projective unitary representation of $\mathbb{Z}_d \times \mathbb{Z}_d$ in $\mathcal{H}$ which we call as the *d-dimensional Weyl representation*. It follows that

$$W_{q,p}^* \mathsf{Q}_j W_{q,p} = \mathsf{Q}_{j-q}, \quad W_{q,p}^* \mathsf{P}_k W_{q,p} = \mathsf{P}_{k-p}, \qquad j, k, q, p \in \mathbb{Z}_d,$$

i.e., Q and P are *Weyl-covariant*.

We denote the set of all Weyl-covariant pairs $(\mathsf{A}, \mathsf{B}) \in \mathbf{Obs}_d \times \mathbf{Obs}_d$, i.e.,

$$(5.5) \qquad W_{q,p}^* \mathsf{A}_j W_{q,p} = \mathsf{A}_{j-q}, \quad W_{q,p}^* \mathsf{B}_k W_{q,p} = \mathsf{B}_{k-p}, \qquad j, k, q, p \in \mathbb{Z}_d,$$

by $\mathbf{Obs}_{d \times d}^W$. Any Weyl-covariant pair (A, B), where A and B are sharp, is unitarily equivalent with the fixed pair (Q, P) in the sense that there is a unitary operator U on $\mathcal{H}$ such that $\mathsf{A}_j = U^* \mathsf{Q}_j U$ and $\mathsf{B}_k = U^* \mathsf{P}_k U$ for all j, $k \in \mathbb{Z}_d$. From now on, we call the pair (Q, P) as the *Weyl pair*.

If $(\mathsf{M}, \mathsf{N}) \in \mathbf{Obs}_{d \times d}^W$, there are probability distributions $\mu = (\mu_j)_{j \in \mathbb{Z}_d}$ and $\nu = (\nu_k)_{k \in \mathbb{Z}_d}$ such that $\mathsf{M} = \mu * \mathsf{Q}$ and $\mathsf{N} = \nu * \mathsf{P}$, i.e.,

$$(5.6) \qquad\qquad \mathsf{M}_j = \sum_{q \in \mathbb{Z}_d} \mu_{j-q} \mathsf{Q}_q, \quad \mathsf{N}_k = \sum_{p \in \mathbb{Z}_d} \nu_{k-p} \mathsf{P}_p$$

for all j, $k \in \mathbb{Z}_d$. Moreover, such a Weyl-covariant pair is jointly measurable if and only if there is a state $\rho \in \mathcal{S}(\mathcal{H})$ such that

$$(5.7) \qquad\qquad \mu_j = \mathrm{tr}[\rho \mathsf{Q}_{-j}], \quad \nu_k = \mathrm{tr}[\rho \mathsf{P}_{-k}], \qquad j, k \in \mathbb{Z}_d.$$

The latter condition can also be written using a purification $\eta \in \mathcal{H} \otimes \mathcal{H}$ of ρ, so that

$$(5.8) \qquad\qquad \mu_j = \langle\eta|(\mathsf{Q}_{-j} \otimes \mathbb{1})\eta\rangle, \quad \nu_k = \langle\eta|(\mathsf{P}_{-k} \otimes \mathbb{1})\eta\rangle, \qquad j, k \in \mathbb{Z}_d.$$

For proofs of these facts about Weyl-covariant pairs, we refer to [5].

The following lemma is useful for evaluating the robustness of incompatibility for any Weyl-covariant pair.

Lemma 1. Let $(\mathsf{A}, \mathsf{B}) \in \mathbf{Obs}_{d \times d}^W$. One has

$$W(\mathsf{A}, \mathsf{B}) = \sup_{(\mathsf{M},\mathsf{N}) \in \mathbf{Obs}_{d \times d}^W} w(\mathsf{A}, \mathsf{B}|\mathsf{M}, \mathsf{N}).$$

Proof. Let us first define a map $\mathbf{Obs}_{d \times d} \ni \mathsf{G} \mapsto \mathsf{G}^W \in \mathbf{Obs}_{d \times d}$ by setting

$$\mathsf{G}_{j,k}^W = \frac{1}{d^2} \sum_{q,p \in \mathbb{Z}_d} W_{q,p}^* \mathsf{G}_{j+q,k+p} W_{q,p}, \quad j, k \in \mathbb{Z}_d.$$

It follows (as one may easily check) that

$$W_{q,p}^* \mathsf{G}_{j,k}^W W_{q,p} = \mathsf{G}_{j-q,k-p}^W$$

for all j, k, q, $p \in \mathbb{Z}_d$. Similarly for any $\mathsf{M} \in \mathbf{Obs}_d$, we define $\mathsf{M}^{W,1}$, $\mathsf{M}^{W,2} \in \mathbf{Obs}_d$ through

$$\mathsf{M}_j^{W,1} \;=\; \frac{1}{d^2} \sum_{q,p \in \mathbb{Z}_d} W_{q,p}^* \mathsf{M}_{j+q} W_{q,p}, \quad j \in \mathbb{Z}_d$$

$$\mathsf{M}_k^{W,2} \;=\; \frac{1}{d^2} \sum_{q,p \in \mathbb{Z}_d} W_{q,p}^* \mathsf{M}_{k+p} W_{q,p}, \quad k \in \mathbb{Z}_d$$

It follows that, when $\mathsf{G} \in \mathbf{Obs}_{d \times d}$ is a joint observable for $(\mathsf{M}, \mathsf{N}) \in \mathbf{JM}_d$, i.e., $\sum_k \mathsf{G}_{j,k} = \mathsf{M}_j$ and $\sum_j \mathsf{G}_{j,k} = \mathsf{N}_k$, then $(\mathsf{M}^{W,1}, \mathsf{N}^{W,2}) \in \mathbf{JM}_d$ and this pair has (among others) the joint observable G^W. Furthermore, $(\mathsf{M}^{W,1}, \mathsf{N}^{W,2}) \in \mathbf{Obs}_{d \times d}^W$ for any M, $\mathsf{N} \in \mathbf{Obs}_d$, and, if $(\mathsf{A}, \mathsf{B}) \in \mathbf{Obs}_{d \times d}^W$, then $\mathsf{A}^{W,1} = \mathsf{A}$ and $\mathsf{B}^{W,2} = \mathsf{B}$.

Suppose now that $t < W(\mathsf{A}, \mathsf{B})$ and let $(\mathsf{M}, \mathsf{N}) \in \mathbf{Obs}_d \times \mathbf{Obs}_d$, and $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \in \mathbf{JM}_d$ be such that

$$t(\mathsf{A}, \mathsf{B}) + (1 - t)(\mathsf{M}, \mathsf{N}) = (\tilde{\mathsf{M}}, \tilde{\mathsf{N}}).$$

Since (A, B) is Weyl covariant, it follows that

$$t(\mathsf{A}, \mathsf{B}) + (1 - t)(\mathsf{M}^{W,1}, \mathsf{N}^{W,2}) = (\tilde{\mathsf{M}}^{W,1}, \tilde{\mathsf{N}}^{W,2}),$$

where $(\tilde{\mathsf{M}}^{W,1}, \tilde{\mathsf{N}}^{W,2}) \in \mathbf{JM}_d$, since, if G is a joint observable for $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}})$, then G^W is a joint observable for $(\tilde{\mathsf{M}}^{W,1}, \tilde{\mathsf{N}}^{W,2})$. Hence, for all $t < W(\mathsf{A}, \mathsf{B})$, we find $(\mathsf{M}, \mathsf{N}) \in \mathbf{Obs}_{d \times d}^W$ and $(\tilde{\mathsf{M}}, \tilde{\mathsf{N}}) \in \mathbf{JM}_d \cap \mathbf{Obs}_{d \times d}^W$ such that

$$t(\mathsf{A}, \mathsf{B}) + (1 - t)(\mathsf{M}, \mathsf{N}) = (\tilde{\mathsf{M}}, \tilde{\mathsf{N}}),$$

and the claim is proven. $\square$

Theorem 4. The robustness of incompatibility for the sharp Weyl pair is given by

$$(5.9) \qquad\qquad\qquad W(\mathsf{Q}, \mathsf{P}) = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{d}}\Big).$$

Proof. Let $t < W(\mathsf{Q}, \mathsf{P})$ and, using Lemma 1, suppose that $(\mathsf{M}, \mathsf{N}) \in \mathbf{Obs}_{d \times d}^W$ and $(\mathsf{A}, \mathsf{B}) \in \mathbf{JM}_d$ are such that

$$(\mathsf{A}, \mathsf{B}) = t(\mathsf{Q}, \mathsf{P}) + (1 - t)(\mathsf{M}, \mathsf{N}).$$

Clearly, $(\mathsf{A}, \mathsf{B}) \in \mathbf{Obs}_{d \times d}^W$. Let μ and ν be probability distributions such that M and N are given by (5.6). Define δ to be the probability distribution having $\delta_0 = 1$ (and, of course, $\delta_j = 0$ for $j \neq 0$). One may write

$$\mathsf{A} = (t\delta + (1 - t)\mu) * \mathsf{Q}, \quad \mathsf{B} = (t\delta + (1 - t)\nu) * \mathsf{P}.$$

Hence, there has to be $\eta \in \mathcal{H} \otimes \mathcal{H}$ such that

$$(5.10) \qquad\qquad \langle \eta | (\mathsf{Q}_{-j} \otimes \mathbb{1})\eta \rangle \;=\; t\delta_j + (1 - t)\mu_j, \quad j \in \mathbb{Z}_d$$

$$(5.11) \qquad\qquad \langle \eta | (\mathsf{P}_{-k} \otimes \mathbb{1})\eta \rangle \;=\; t\delta_k + (1 - t)\nu_k, \quad k \in \mathbb{Z}_d.$$

We may write $\eta = \sum_j \varphi_j \otimes \zeta_j = \sum_k \psi_k \otimes \xi_k$ for some ζ_j, $\xi_k \in \mathcal{H}$, j, $k \in \mathbb{Z}_d$. Following the procedure carried out in the proof of [5, Lemma 1], one obtains the (tight) inequalities

$$(5.12) \qquad\quad \sqrt{t(1 - \nu_0) + \nu_0} \;\leq\; \frac{1}{\sqrt{d}}\Big(\sqrt{t(1 - \mu_0) + \mu_0} + \alpha\sqrt{1 - t}\Big)$$

$$(5.13) \qquad\quad \sqrt{t(1 - \mu_0) + \mu_0} \;\leq\; \frac{1}{\sqrt{d}}\Big(\sqrt{t(1 - \nu_0) + \nu_0} + \beta\sqrt{1 - t}\Big),$$

where $\alpha = \sum_{j \neq 0} \sqrt{\mu_j}$ and $\beta = \sum_{k \neq 0} \sqrt{\nu_k}$. It is easy to see that $\alpha \leq \sqrt{d-1}\sqrt{1-\mu_0}$ and $\beta \leq \sqrt{d-1}\sqrt{1-\nu_0}$, and these bounds are reached when $\mu_j = (1-\mu_0)/(d-1)$ and $\nu_k = (1-\nu_0)/(d-1)$ for all j, $k \neq 0$. Solving from (5.12)-(5.13), one obtains the following inequalities:

$$t \leq \frac{(\alpha + \sqrt{d}\beta)^2 - (d-1)^2\mu_0}{(\alpha + \sqrt{d}\beta)^2 + (d-1)^2(1-\mu_0)} \leq \frac{(d-1)\left(\sqrt{1-\mu_0} + \sqrt{d(1-\nu_0)}\right)^2 - (d-1)^2\mu_0}{(d-1)\left(\sqrt{1-\mu_0} + \sqrt{d(1-\nu_0)}\right)^2 + (d-1)^2(1-\mu_0)}.$$

It is easy to see that as one lets μ_0, $\nu_0 \downarrow 0$, the latter bound increases and setting $\mu_0 = \nu_0 = 0$, one obtains $W(\mathsf{Q}, \mathsf{P}) \leq \frac{1}{2}(1 + 1/\sqrt{d})$.

It remains to be shown that the bound obtained above is reachable. Setting $t = \frac{1}{2}(1 + 1/\sqrt{d})$,

$$\eta = \sqrt{\frac{\sqrt{d}}{2(\sqrt{d} + 1)}}(\varphi_0 + \psi_0) \otimes \xi$$

for any unit vector $\xi \in \mathcal{H}$, $\mu_0 = \nu_0 = 0$, and $\mu_j = \nu_j = 1/(d-1)$, for $j \neq 0$, one finds that Equations (5.10)-(5.11) hold (and in (5.12)-(5.13) the inequalities can both be replaced by equalities). Hence, the claim is proven. $\square$

5.2. **Robustness of incompatibility for a pair of decodable channels.** Let us fix a finite-dimensional Hilbert space $\mathcal{H}$, $\dim \mathcal{H} = d$. We denote $\mathbf{Ch}(\mathcal{H}, \mathcal{H}) = \mathbf{Ch}_d$ and $\mathbf{Ch}(\mathcal{H}, \mathcal{H} \otimes \mathcal{H}) = \mathbf{Ch}_{d \times d}$. The set of compatible pairs within $\mathbf{Ch}_d \times \mathbf{Ch}_d$ is denoted by $\mathbf{Comp}_d$. Moreover, id stands for the identity channel $\mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$, i.e., $\mathrm{id}(\rho) = \rho$ for all $\rho \in \mathcal{S}(\mathcal{H})$. Clearly, the dual id^* is the identity map on $\mathcal{L}(\mathcal{H})$ which we denote by id as well.

We fix an orthonormal basis $\{|n\rangle\}_{n=1}^d$ for the duration of this subsection and introduce the rank-1 operators $\Omega_d = \sum_{m,n=1}^d |mm\rangle\langle nn| \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ and $\Omega_{d \otimes d} = \sum_{j,k,m,n=1}^d |jkjk\rangle\langle mnmn| \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$, where $|m_1 \cdots m_n\rangle := |m_1\rangle \otimes \cdots \otimes |m_n\rangle$. For any $\mathcal{E} \in \mathbf{Ch}_d$ (respectively $\mathcal{F} \in \mathbf{Ch}_{d \times d}$) we define the *Choi operator* $M(\mathcal{E}) = (\mathcal{E}^* \otimes \mathrm{id})(\Omega_d) \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ (respectively $M(\mathcal{F}) = (\mathcal{F}^* \otimes \mathrm{id} \otimes \mathrm{id})(\Omega_{d \times d}) \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$). We denote the transpose of $B \in \mathcal{L}(\mathcal{H})$ with respect to the fixed basis by B^T and denote $\overline{B} := B^{T*}$. Furthermore, we denote the partial transpose restricted to the subsystems 2 and 3 of $C \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$ by C^Γ, i.e., when $C_1, C_2, C_3 \in \mathcal{L}(\mathcal{H})$, we have $(C_1 \otimes C_2 \otimes C_3)^\Gamma = C_1 \otimes C_2^T \otimes C_3^T$.

If $\mathcal{E} \in \mathbf{Ch}_d$ (respectively $\mathcal{F} \in \mathbf{Ch}_{d \times d}$) is such that $\mathcal{E}(U\rho U^*) = U\mathcal{E}(\rho)U^*$ (respectively $\mathcal{F}(U\rho U^*) = (U \otimes U)\mathcal{F}(\rho)(U \otimes U)^*$) for all $\rho \in \mathcal{S}(\mathcal{H})$ and all unitary $U \in \mathcal{L}(\mathcal{H})$, we say that $\mathcal{E}$ (respectively $\mathcal{F}$) is *fully covariant* and denote $\mathcal{E} \in \mathbf{Cov}_d$ (respectively $\mathcal{F} \in \mathbf{Cov}_{d \times d}$). Denote by dU the normalized Haar measure of the (compact) unitary group $U(d)$. We may define the map $\mathbf{Ch}_d \ni \mathcal{E} \mapsto \mathcal{E}_{av} \in \mathbf{Cov}_d$ by setting

$$\mathcal{E}_{av}(\rho) = \int_{U(d)} U^*\mathcal{E}(U\rho U^*)U \, dU, \qquad \rho \in \mathcal{S}(\mathcal{H}).$$

Likewise, one can set up a map $\mathbf{Ch}_{d \times d} \ni \mathcal{F} \mapsto \mathcal{F}_{av} \in \mathbf{Cov}_{d \times d}$ through

$$(5.14) \qquad \mathcal{F}_{av}(\rho) = \int_{U(d)} (U \otimes U)^*\mathcal{F}(U\rho U^*)(U \otimes U) \, dU, \qquad \rho \in \mathcal{S}(\mathcal{H}).$$

The sets $\mathbf{Cov}_d$ and $\mathbf{Cov}_{d \times d}$ coincide with the sets of the fixed points of these maps.

Suppose that $M \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$ is a positive operator whose partial trace over the subsystems 2 and 3 coincides with $\mathbb{1}_{\mathcal{H}}$, i.e., M is a Choi operator of a channel $\mathcal{F} \in \mathbf{Ch}_{d \times d}$. We may define the operator $M_{av} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$ through

$$(5.15) \qquad M_{av} = \int_{U(d)} (U \otimes \overline{U} \otimes \overline{U})M(U \otimes \overline{U} \otimes \overline{U})^* \, dU.$$

We have that $(U \otimes \overline{U} \otimes \overline{U})M = M(U \otimes \overline{U} \otimes \overline{U})$ for all $U \in U(d)$ if and only if $M = M_{av}$. It is straightforward to check that

$$M(\mathcal{F}_{av}) = M(\mathcal{F})_{av}, \qquad \mathcal{F} \in \mathbf{Ch}_{d \times d}.$$

Lemma 2. Let $\mathcal{E} \in \mathbf{Cov}_d$. The robustness of self-incompatibility for $\mathcal{E}$ is given by

$$W(\mathcal{E}) = \sup_{\mathcal{C} \in \mathbf{Cov}_d} w(\mathcal{E}, \mathcal{E} | \mathcal{C}, \mathcal{C}).$$

Proof. Let $t < W(\mathcal{E})$ and suppose that $(\mathcal{A}, \mathcal{B}) \in \mathbf{Ch}_d \times \mathbf{Ch}_d$ and $(\mathcal{A}', \mathcal{B}') \in \mathbf{Comp}_d$ are such that

$$t(\mathcal{E}, \mathcal{E}) + (1 - t)(\mathcal{A}, \mathcal{B}) = (\mathcal{A}', \mathcal{B}')$$

It follows that $(\mathcal{A}'_{av}, \mathcal{B}'_{av}) \in \mathbf{Comp}_d$ as well, since if $\mathcal{F}$ is a joint channel for $(\mathcal{A}', \mathcal{B}')$, it is immediate that $\mathcal{F}_{av}$ is a joint channel for $(\mathcal{A}'_{av}, \mathcal{B}'_{av})$. Hence,

$$t(\mathcal{E}, \mathcal{E}) + (1 - t)(\mathcal{A}_{av}, \mathcal{B}_{av}) = (\mathcal{A}'_{av}, \mathcal{B}'_{av}).$$

Denote $\mathcal{C} = \frac{1}{2}(\mathcal{A}_{av} + \mathcal{B}_{av})$ and $\mathcal{C}' = \frac{1}{2}(\mathcal{A}'_{av} + \mathcal{B}'_{av})$. Again it easily follows

$$t(\mathcal{E}, \mathcal{E}) + (1 - t)(\mathcal{C}, \mathcal{C}) = (\mathcal{C}', \mathcal{C}').$$

Denote by $F \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ the flip operator, $F(\varphi \otimes \psi) = \psi \otimes \varphi$ for all $\varphi, \psi \in \mathcal{H}$. The pair $(\mathcal{C}', \mathcal{C}')$ is compatible, since if M is the Choi operator of a joint channel of $(\mathcal{A}'_{av}, \mathcal{B}'_{av})$, then $\frac{1}{2}\big(M + (\mathbb{1} \otimes F)M(\mathbb{1} \otimes F)\big)$ is the Choi operator for a joint channel for $(\mathcal{C}', \mathcal{C}')$. $\qquad\square$

Since id is fully covariant, the preceding lemma restricts the problem of evaluating the robustness of self-incompatibility of id (quite considerably, as we will see). According to Lemma 2 (and its proof), $W(\mathrm{id})$ is simply the supremum of those $t \in [0, 1]$ such that $t\,\mathrm{id} + (1 - t)\mathcal{E}$ is self-compatible for some $\mathcal{E} \in \mathbf{Cov}_d$. Next, we determine the set of self-compatible fully covariant channels which essentially resolves the problem of determining $W(\mathrm{id})$.

Let us fix a self-compatible $\mathcal{A} \in \mathbf{Cov}_d$ and a joint channel $\mathcal{F} \in \mathbf{Ch}_{d \times d}$ for the pair $(\mathcal{A}, \mathcal{A})$. Since $\mathcal{A}$ is fully covariant, the channel $\mathcal{F}_{av}$ still has the same marginals. Let $M \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$ be the Choi operator of $\mathcal{F}$ (with respect to our fixed basis). Hence, M_{av} is the Choi operator of $\mathcal{F}_{av}$. From (5.15) it follows that

$$M_{av} = \left(\int (U \otimes U \otimes U)M^{\Gamma}(U \otimes U \otimes U)^{*}\, dU \right)^{\Gamma}.$$

This means that $(U \otimes U \otimes U)M_{av}^{\Gamma} = M_{av}^{\Gamma}(U \otimes U \otimes U)$ for all $U \in U(d)$, i.e., M_{av}^{Γ} is $U \otimes U \otimes U$-invariant.

For any permutation π of three elements, denote by $V_{\pi} \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H} \otimes \mathcal{H})$ the unitary operator defined through

$$V_{\pi}(\varphi_1 \otimes \varphi_2 \otimes \varphi_3) = \varphi_{\pi^{-1}(1)} \otimes \varphi_{\pi^{-1}(2)} \otimes \varphi_{\pi^{-1}(3)}$$

for all $\varphi_1, \varphi_2, \varphi_3 \in \mathcal{H}$. A well-known result from Weyl states that any $U \otimes U \otimes U$-invariant operator is a linear combination of the permutation operators V_{π} from which it follows immediately that our averaged Choi operator can be expressed as a linear combination

$$M_{av} = \sum_{\pi} \lambda_{\pi} V_{\pi}^{\Gamma},$$

where λ_{π} are complex numbers. However, we must make sure that this linear combination is a positive operator whose partial trace over the subsystems 2 and 3 is $\mathbb{1}_{\mathcal{H}}$. To this end, we must express the linear combination in a more revealing form.

As the commutant of the set $U \otimes \overline{U} \otimes \overline{U}$, $U \in U(d)$, the operator system spanned by the six operators V_π^Γ is a 6-dimensional algebra with the exception in case $d = 2$, when the algebra is 5-dimensional. In [7], this algebra was shown to have the basis consisting of the operators

$$S_\pm = \frac{1}{2}\big(\mathbb{1} \pm V_{(23)}^\Gamma - \frac{1}{d \pm 1}(V_{(12)}^\Gamma + V_{(13)}^\Gamma \pm V_{(123)}^\Gamma \pm V_{(132)}^\Gamma)\big),$$

$$S_0 = \frac{1}{d^2 - 1}\big(d(V_{(12)}^\Gamma + V_{(13)}^\Gamma) - (V_{(123)}^\Gamma + V_{(132)}^\Gamma)\big),$$

$$S_1 = \frac{1}{d^2 - 1}\big(d(V_{(123)}^\Gamma + V_{(132)}^\Gamma) - (V_{(12)}^\Gamma + V_{(13)}^\Gamma)\big),$$

$$S_2 = \frac{1}{\sqrt{d^2 - 1}}(V_{(12)}^\Gamma - V_{(13)}^\Gamma),$$

$$S_3 = \frac{i}{\sqrt{d^2 - 1}}(V_{(123)}^\Gamma - V_{(132)}^\Gamma),$$

where $S_\pm$, S_0 are mutually orthogonal projections summing up to $\mathbb{1}$ and S_1, S_2, and S_3 are selfadjoint operators supported on the eigenspace of S_0 that are interrelated in the same way as the Pauli matrices. Hence, formally $S_j S_\pm = S_\pm S_j = 0$ and $S_j^2 = S_0$ for all $j = 0, 1, 2, 3$, and $S_1 S_2 = i S_3$ with cyclic permutations. Note that, in the case $d = 2$, $S_- = 0$. It follows that a linear combination $\mu_+ S_+ + \mu_- S_- + \mu_0 S_0 + \mu_1 S_1 + \mu_2 S_2 + \mu_3 S_3$ is positive if and only if the multipliers are real, $\mu_\pm$, $\mu_0 \geq 0$, and $\mu_1^2 + \mu_2^2 + \mu_3^2 \leq \mu_0^2$.

Let us now impose our additional symmetry condition, i.e., the marginals of $\mathcal{F}_{av}$ must coincide or, equivalently, $V_{(23)} M_{av} = M_{av} V_{(23)}$; note that $V_{(23)}^\Gamma = V_{(23)}$. One easily sees that this requirement necessitates that the multipliers of S_2 and S_3 in M be zero. Moreover, through direct calculation, one finds that for the partial traces over the subsystems 2 and 3,

$$\begin{array}{llll}
\mathrm{tr}_{23}[S_+] &=& \frac{1}{2}(d-1)(d+2)\mathbb{1}_\mathcal{H}, & \mathrm{tr}_{23}[S_-] = \frac{1}{2}(d+1)(d-2)\mathbb{1}_\mathcal{H}, \\
\mathrm{tr}_{23}[S_0] &=& 2\mathbb{1}_\mathcal{H}, & \mathrm{tr}_{23}[S_1] = 0.
\end{array}$$

Putting all this together, one finds that M_{av} is of the form $M(t_+, t_-, t_0, t_1)$,

$$M(t_+, t_-, t_0, t_1) = \frac{2}{(d-1)(d+2)}t_+ S_+ + \frac{2}{(d+1)(d-2)}t_- S_- + \frac{1}{2}(t_0 S_0 + t_1 S_1),$$

where $t_\pm, t_0 \geq 0$, $t_+ + t_- + t_0 = 1$ and $t_1 \in [-t_0, t_0]$.

The set of the Choi operators $M(t_+, t_-, t_0, t_1)$ is a tetrahedron with the extreme points

$$M_\pm = \frac{2}{(d \mp 1)(d \pm 2)}S_\pm,$$

$$\tilde{M}_\pm = \frac{1}{2}(S_0 \pm S_1),$$

and the partial traces over system 2 (or, equivalently, over system 3) of these are

$$\mathrm{tr}_2[M_\pm] = \frac{1}{d^2 - 1}(d\mathbb{1}_{\mathcal{H} \otimes \mathcal{H}} - \Omega_d),$$

$$\mathrm{tr}_2[\tilde{M}_\pm] = \frac{1}{2(d \pm 1)}\big(\mathbb{1}_{\mathcal{H} \otimes \mathcal{H}} + (d \pm 2)\Omega_d\big).$$

Hence, the (coinciding) marginals of the channels corresponding to $M_\pm$ are $(d^2 - 1)^{-1}(d^2 \mathcal{T} - \mathrm{id})$ and the (coinciding) marginals of the channels corresponding to $\tilde{M}_\pm$ are $\big(2(d \pm 1)\big)^{-1}\big(d\mathcal{T} + (d \pm 2)\mathrm{id}\big)$, where $\mathcal{T} : \mathcal{S}(\mathcal{H}) \to \mathcal{S}(\mathcal{H})$ is the constant channel $\mathcal{T}(\rho) = d^{-1}\mathbb{1}_\mathcal{H}$. Thus, the set of self-compatible fully covariant channels consists of the elements $\lambda\mathcal{T} + (1 - \lambda)\mathrm{id}$ where $\frac{d}{2(d+1)} \leq$

$\lambda \leq \frac{d^2}{d^2-1}$. It follows that $W(\mathrm{id})$ is the supremum of the $t \in [0,1]$ such that

$$\frac{1}{2(1+d)}\big(d\mathfrak{T} + (d+2)\mathrm{id}\big) - t\,\mathrm{id}$$

is completely positive or, using the Choi operators, $\mathrm{tr}_2[\tilde{M}_+] - t\Omega_d \geq 0$. One easily finds that this condition is satisfied if and only if $t \leq \frac{1}{2}(1 + 1/d)$. Thus, $W(\mathrm{id}) = \frac{1}{2}(1 + 1/d)$. According to the discussion following Theorem 2, a pair $(\mathcal{V}, \mathcal{W}) \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1) \times \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$ of decodable channels has the same robustness of incompatibility as the pair of identity channels which is the minimum of the robustness measure, and hence:

Theorem 5. Any pair $(\mathcal{V}, \mathcal{W}) \in \mathbf{Ch}(\mathcal{H}, \mathcal{K}_1) \times \mathbf{Ch}(\mathcal{H}, \mathcal{K}_2)$ of decodable channels minimizes W amongst the channel pairs with the input space $\mathcal{H}$, and this minimum value is

$$W(\mathcal{V}, \mathcal{W}) = W(\mathrm{id}) = \frac{1}{2}\Big(1 + \frac{1}{d}\Big),$$

where $d = \dim \mathcal{H}$ and id is the identity channel on $\mathcal{S}(\mathcal{H})$.

According to the discussion preceding Theorem 5, the self-compatible fully covariant channel $\mathcal{A}$ that has the optimality property

(5.16) $$\mathcal{A} = \frac{1}{2}\Big(1 + \frac{1}{d}\Big)\,\mathrm{id} + \frac{1}{2}\Big(1 - \frac{1}{d}\Big)\mathcal{E}$$

for some $\mathcal{E} \in \mathbf{Ch}_d$ is associated with the Choi operator $\mathrm{tr}_2[\tilde{M}_+]$ and is thus given by

$$\mathcal{A} = \frac{d+2}{2(d+1)}\mathrm{id} + \frac{d}{2(d+1)}\mathfrak{T}.$$

One joint channel for the optimal pair $(\mathcal{A}, \mathcal{A})$ is thus the optimal universal cloner $\rho \mapsto 2(d+1)^{-1}S(\rho \otimes \mathbb{1}_{\mathcal{H}})S$, where $S \in \mathcal{L}(\mathcal{H} \otimes \mathcal{H})$ is the orthogonal projector onto the symmetric subspace (the 1-eigenspace of the flip operator) of $\mathcal{H} \otimes \mathcal{H}$. In fact, the joint channel associated with the Choi operator $\tilde{M}_+$ is exactly this optimal cloner. The Choi operator associated with the channel $\mathcal{E}$ in the decomposition (5.16) is $\mathrm{tr}_2[M_+] = \mathrm{tr}_2[M_-]$, so that

$$\mathcal{E} = -\frac{1}{d^2-1}\mathrm{id} + \frac{d^2}{d^2-1}\mathfrak{T}.$$

Especially $\mathcal{E}$ is self-compatible so that, in fact

$$W(\mathrm{id}) = w(\mathrm{id}, \mathrm{id}) = \frac{1}{2}\Big(1 + \frac{1}{d}\Big).$$

The same holds for any pair $(\mathcal{V}, \mathcal{W})$ of decodable channels. Moreover the pair

$$\frac{1}{2}\Big(1 + \frac{1}{d}\Big)(\mathcal{V}, \mathcal{W}) + \frac{1}{2}\Big(1 - \frac{1}{d}\Big)(\mathcal{E}_\mathcal{V}, \mathcal{E}_\mathcal{W}) = \frac{d+2}{2(d+1)}(\mathcal{V}, \mathcal{W}) + \frac{d}{2(d+1)}(\mathfrak{T}, \mathfrak{T})$$

is compatible, where

$$(\mathcal{E}_\mathcal{V}, \mathcal{E}_\mathcal{W}) = -\frac{1}{d^2-1}(\mathcal{V}, \mathcal{W}) + \frac{d^2}{d^2-1}(\mathfrak{T}, \mathfrak{T}).$$

5.3. **Robustness of incompatibility for a von Neumann observable and a decodable channel.** In this subsection, we calculate the robustness of incompatibility for a von Neumann observable, i.e. a rank-1 sharp observable (PVM), and a decodable channel. We start with the (slightly) simpler case where the decodable channel is simply the identity channel from which the more general case follows according to Section 4.2.

For the remainder of this subsection, we fix a d-dimensional Hilbert space ($d < \infty$), and we fix an orthonormal basis $\{|n\rangle\}_{n=0}^{d-1} \subset \mathcal{H}$; we treat the index set $\{0, 1, \ldots, d-1\}$ as the cyclic group $\mathbb{Z}_d$ and all sums and differences of these indices are considered cyclic. For simplicity, we denote the set of observables on the power set of $\mathbb{Z}_d$ and operating in $\mathcal{H}$ by $\mathbf{Obs}_d$ and the set of channels $\mathbf{Ch}(\mathcal{H}, \mathcal{H})$ by $\mathbf{Ch}_d$. We denote the set of instruments on the power set of $\mathbb{Z}_d$ and operating within $\mathcal{S}(\mathcal{H})$ by $\mathbf{Ins}_d$, i.e., instruments in $\mathbf{Ins}_d$ are the possible joint instruments for pairs $(\mathsf{M}, \mathcal{E}) \in \mathbf{Obs}_d \times \mathbf{Ch}_d$. We treat any $\Gamma \in \mathbf{Ins}_d$ as a sequence $(\Gamma_j)_{j \in \mathbb{Z}_d}$ of operations summing up to a trace-preserving operation.

For any $q, p \in \mathbb{Z}_d$, we define the operators $U_q, V_p, W_{q,p} \in \mathcal{L}(\mathcal{H})$ in the same way as in Equations (5.2)-(5.4) with the basis $\{\varphi_j\}_{j \in \mathbb{Z}_d}$ replaced with the basis $\{|n\rangle\}_{n \in \mathbb{Z}_d}$, so that, e.g., $U_q|n\rangle = |n+q\rangle$ for all $q, n \in \mathbb{Z}_d$. We denote the set of those $\mathsf{M} \in \mathbf{Obs}_d$ such that $W_{q,p}^* \mathsf{M}_j W_{q,p} = \mathsf{M}_{j-q}$ for all $q, p, j \in \mathbb{Z}_d$ by $\mathbf{Obs}_d^{W,1}$. Furthermore, we denote by $\mathbf{Ch}_d^{W,2}$ the set of those $\mathcal{E} \in \mathbf{Ch}_d$ such that $\mathcal{E}(W_{q,p}\rho W_{q,p}^*) = W_{q,p}\mathcal{E}(\rho)W_{q,p}^*$ for all $q, p \in \mathbb{Z}_d$ and all $\rho \in \mathcal{S}(\mathcal{H})$. Finally, we denote the set of those $\Gamma \in \mathbf{Ins}_d$ such that $\Gamma_{j-q}(W_{q,p}\rho W_{q,p}^*) = W_{q,p}\Gamma_j(\rho)W_{q,p}^*$ for all $j, q, p \in \mathbb{Z}_d$ and all $\rho \in \mathcal{S}(\mathcal{H})$ by $\mathbf{Ins}_d^W$.

The elements in the sets $\mathbf{Obs}_d^{W,1}$, $\mathbf{Ch}_d^{W,2}$, and $\mathbf{Ins}_d^W$ have a simple structure: For any $\mathsf{M} \in \mathbf{Obs}_d^{W,1}$, there is a positive operator $C \in \mathcal{L}(\mathcal{H})$ such that $V_p C = C V_p$ for all $p \in \mathbb{Z}_d$, $\sum_j U_j C U_j^* = \mathbb{1}_{\mathcal{H}}$ (so that, especially, $\mathrm{tr}[C] = 1$), and

$$(5.17) \qquad \mathsf{M}_j = U_j C U_j^*, \qquad j \in \mathbb{Z}_d.$$

For each $\Gamma \in \mathbf{Ins}_d^W$, there is an operation $\mathcal{D} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$ such that $\mathcal{D}(V_p\rho V_p^*) = V_p\mathcal{D}(\rho)V_p^*$ for all $p \in \mathbb{Z}_d$ and $\rho \in \mathcal{S}(\mathcal{H})$, and $\sum_j U_j \mathcal{D}^*(\mathbb{1}_{\mathcal{H}})U_j^* = \mathbb{1}_{\mathcal{H}}$, and

$$(5.18) \qquad \Gamma_j(\rho) = U_j \mathcal{D}(U_j^* \rho U_j) U_j^*, \qquad j \in \mathbb{Z}_d, \quad \rho \in \mathcal{S}(\mathcal{H}).$$

For the characterizations given above for W-covariant observables and instruments, see [15, Section III]; the conjecture presented in the reference certainly holds in our discrete case. For any covariant channel $\mathcal{E} \in \mathbf{Ch}_d^{W,2}$ there is a positive kernel $(q, p) \mapsto \Phi_{q,p} \in \mathbb{C}$ such that [16]

$$(5.19) \qquad \mathcal{E}^*(W_{q,p}) = \Phi_{q,p} W_{q,p}, \qquad q, p \in \mathbb{Z}_d.$$

Since the operators $W_{q,p}$, $q, p \in \mathbb{Z}_d$, span the whole of $\mathcal{L}(\mathcal{H})$, the kernel Φ completely characterizes the covariant channel. Positivity of the kernel Φ means that the Fourier transform of the kernel is positive, i.e., for any $j, k \in \mathbb{Z}_d$,

$$(5.20) \qquad \hat{\Phi}_{j,k} := \frac{1}{d} \sum_{q,p \in \mathbb{Z}_d} \overline{\langle q, k\rangle}\langle j, p\rangle \Phi_{q,p} \geq 0.$$

Moreover, when $\mathcal{E}$ is defined as in (5.19), then

$$\mathcal{E}(\rho) = \frac{1}{d} \sum_{j,k \in \mathbb{Z}_d} \hat{\Phi}_{j,k} W_{j,k}\rho W_{j,k}^*, \qquad \rho \in \mathcal{S}(\mathcal{H}).$$

When the channel $\mathcal{E}$ arises from a covariant instrument like that in Equation (5.18), it follows from straight-forward calculation (utilizing the fact that, for any $B \in \mathcal{L}(\mathcal{H})$, one has $\sum_{q,p} W_{q,p} B W_{q,p}^* = d\,\mathrm{tr}[B]\mathbb{1}$) that the kernel Φ associated with $\mathcal{E}$ is of the form

$$(5.21) \qquad \Phi_{q,p} = \mathrm{tr}[W_{q,p}^* \mathcal{D}^*(W_{q,p})], \qquad q, p \in \mathbb{Z}_d.$$

As earlier, we have the covariantization maps $\mathbf{Obs}_d \ni \mathsf{M} \mapsto \mathsf{M}^{W,1} \in \mathbf{Obs}_d^{W,1}$, $\mathbf{Ch}_d \ni \mathcal{E} \mapsto \mathcal{E}^{W,2} \in \mathbf{Ch}_d^{W,2}$, and $\mathbf{Ins}_d \ni \Gamma \mapsto \Gamma^W \in \mathbf{Ins}_d^W$ having the covariant devices as their fixed points. Especially,

$$\Gamma_j^W(\rho) = \frac{1}{d^2} \sum_{q,p \in \mathbb{Z}_d} W_{q,p}^* \Gamma_{j-q}(W_{q,p} \rho W_{q,p}^*) W_{q,p}$$

for any $\Gamma \in \mathbf{Ins}_d$, $j \in \mathbb{Z}_d$, and $\rho \in \mathcal{S}(\mathcal{H})$. Moreover, by similar arguments as earlier, one can show that, if a pair $(\mathsf{M}, \mathcal{E}) \in \mathbf{Obs}_d \times \mathbf{Ch}_d$ is compatible, so is $(\mathsf{M}^{W,1}, \mathcal{E}^{W,2})$, and if a pair $(\mathsf{M}, \mathcal{E}) \in \mathbf{Obs}_d^{W,1} \times \mathbf{Ch}_d^{W,2}$ is compatible, it has a joint instrument in $\mathbf{Ins}_d^W$. As in preceding analyses, it follows:

Lemma 3. Suppose that $(\mathsf{A}, \mathcal{A}) \in \mathbf{Obs}_d^{W,1} \times \mathbf{Ch}_d^{W,2}$. One has

$$W(\mathsf{A}, \mathcal{A}) = \sup\{w(\mathsf{A}, \mathcal{A}|\mathsf{B}, \mathcal{B}) \,|\, \mathsf{B} \in \mathbf{Obs}_d^{W,1}, \ \mathcal{B} \in \mathbf{Ch}_d^{W,2}\}.$$

In what follows, we study the robustness of incompatibility of the pair $(\mathsf{A}, \mathrm{id})$, where $\mathsf{A}_n = |n\rangle\langle n|$ for all $n \in \mathbb{Z}_d$. Evidently, $\mathsf{A} \in \mathbf{Obs}_d^{W,1}$ and $\mathrm{id} \in \mathbf{Ch}_d^{W,2}$, but this pair is not compatible. According to the preceding lemma, there is a pair $(\mathsf{B}, \mathcal{B}) \in \mathbf{Obs}_d^{W,1} \times \mathbf{Ch}_d^{W,2}$ such that the pair

$$\big(W\mathsf{A} + (1-W)\mathsf{B}, W\mathrm{id} + (1-W)\mathcal{B}\big),$$

with $W := W(\mathsf{A}, \mathrm{id})$, is compatible, and hence has a joint instrument in $\mathbf{Ins}_d^W$. We start determining the value of $W(\mathsf{A}, \mathrm{id})$ first by giving a characterization for the instruments of $\mathbf{Ins}_d^W$.

Lemma 4. For any $\Gamma \in \mathbf{Ins}_d^W$ there is an indexed set $\alpha = (\alpha_{r,s}^n)_{n,r,s \in \mathbb{Z}_d}$ of complex numbers such that, for all $n \in \mathbb{Z}_d$, the matrix $(\alpha_{r,s}^n)_{r,s \in \mathbb{Z}_d}$ is positive, $\sum_{n,r} \alpha_{r,r}^n = 1$, and defining $\mathcal{D} : \mathcal{L}(\mathcal{H}) \to \mathcal{L}(\mathcal{H})$,

$$(5.22) \qquad \mathcal{D}(A) = \sum_{n,r,s \in \mathbb{Z}_d} \alpha_{r,s}^n \langle n+r|A|n+s\rangle |r\rangle\langle s|, \qquad A \in \mathcal{L}(\mathcal{H}),$$

Γ is defined as in Equation (5.18).

Proof. Suppose that $\Gamma \in \mathbf{Ins}_d^W$ is associated with an operation $\mathcal{D}$ that is covariant with respect to the representation $p \mapsto V_p$ (i.e., is V-covariant), $\mathrm{tr}[\mathcal{D}^*(\mathbb{1})] = 1$, and Γ is defined as in (5.18). The V-covariance means that, for the Choi operator $M = (\mathcal{D}^* \otimes \mathrm{id})(\Omega_d)$, where $\Omega_d = \sum_{m,n} |m,m\rangle\langle n,n|$, we have

$$(V_p \otimes \overline{V}_p)M = M(V_p \otimes \overline{V}_p), \qquad p \in \mathbb{Z}_d,$$

where $\overline{V}_p = V_p^{*T}$ with the transpose defined with respect to the basis $\{|n\rangle\}_n$. Since the representation $p \mapsto V_p \otimes \overline{V}_p$ is associated with the spectral measure $q \mapsto \mathsf{R}_q$,

$$\mathsf{R}_q = \sum_{r \in \mathbb{Z}_d} |q+r, r\rangle\langle q+r, r|, \qquad q \in \mathbb{Z}_d,$$

i.e., $V_p \otimes \overline{V}_p = \sum_q \langle q, p \rangle \mathsf{R}_q$, it follows that any operator that commutes with this representation has to be a linear combination of the operators $A_{n,r,s} = |n+r, r\rangle\langle n+s, s|$, $n, r, s \in \mathbb{Z}_d$. Writing $M = \sum_{n,r,s} \alpha_{r,s}^n A_{n,r,s}$, one finds that $\mathcal{D}$ is defined as in (5.22). The positivity of M is equivalent with $\mathsf{R}_n M \mathsf{R}_n \geq 0$ for all n which, in turn, is equivalent with $(\alpha_{r,s}^n)_{r,s} \geq 0$ for all n. Direct calculation shows that $\sum_j U_j \mathcal{D}^*(\mathbb{1}) U_j^* = \mathbb{1}$ is equivalent with $\sum_{n,r} \alpha_{r,r}^n = 1$. $\qquad\square$

Let us pick a compatible pair $(\mathsf{M}, \mathcal{E}) \in \mathbf{Obs}_d^{W,1} \times \mathbf{Ch}_d^{W,2}$ that has a joint instrument $\Gamma \in \mathbf{Ins}_d^W$ that is associated with α like that in Lemma 4. Through simple calculations one finds that

the trace-1 positive operator C associated with M according to (5.17) and the positive kernel Φ associated with $\mathcal{E}$ according to (5.19) are given by

$$(5.23) \qquad C = \sum_{n,r\in\mathbb{Z}_d} \alpha_{r,r}^n |n+r\rangle\langle n+r|, \qquad \Phi_{q,p} = \sum_{n,r\in\mathbb{Z}_d} \overline{\langle n,p\rangle} \alpha_{r,r-q}^n, \qquad q,\, p \in \mathbb{Z}_d.$$

In the case of our special incompatible pair $(\mathsf{A}, \mathrm{id})$, the trace-1 positive operator associated with A is $|0\rangle\langle 0|$ and the positive kernel associated with the identity channel is the constant kernel 1. It follows that, if we require the compatible pair like that above to be a convex combination of the form

$$(\mathsf{M}, \mathcal{E}) = t(\mathsf{A}, \mathrm{id}) + (1-t)(\mathsf{B}, \mathcal{B})$$

with some $t \in [0,1]$ and some (not necessarily compatible) pair $(\mathsf{B}, \mathcal{B}) \in \mathbf{Obs}_d \times \mathbf{Ch}_d$, then we must have $D = C - t|0\rangle\langle 0| \geq 0$ and the kernel $\Psi = \Phi - t$ has to be positive, where C and Φ are defined as in (5.23). Since C is diagonalized in the basis $\{|n\rangle\}_n$, it is a straight-forward check that the first condition is equivalent with

$$(5.24) \qquad t \leq \sum_{n\in\mathbb{Z}_d} \alpha_{-n,-n}^n =: w_1(\alpha).$$

Using the characterization of (5.20) for the positivity of the kernel Ψ, the latter condition is equivalent with

$$(5.25) \qquad t \leq \frac{1}{d} \sum_{r,s\in\mathbb{Z}_d} \alpha_{r,s}^0 =: w_2(\alpha).$$

The robustness $W(\mathsf{A}, \mathrm{id})$ is simply the supremum of $\min\{w_1(\alpha), w_2(\alpha)\}$ over all those $\alpha = (\alpha_{r,s}^n)_{n,r,s}$ such that $(\alpha_{r,s}^n)_{r,s} \geq 0$ for all $n \in \mathbb{Z}_d$ and $\sum_{n,r} \alpha_{r,r}^n = 1$.

We may simplify the optimization task presented above by a couple of observations: First, we note that, given α, the elements $\alpha_{r,s}^n$ where $n \neq 0$, $r \neq -n$, and $s \neq -n$ can be assumed to be zero; this assumption does not affect the value of w_2 and it can only increase the value of w_1. The elements $\alpha_{-n,-n}^n$ are, of course, non-negative. Second, the property $\sum_{n,r} \alpha_{r,r}^n = 1$ and the positivity requirements are not violated if we replace the α, where the elements $\alpha_{r,s}^n$ with $n \neq 0$, $r \neq -n$, and $s \neq -n$ are zero, with $\tilde{\alpha}$, where $\tilde{\alpha}_{0,0}^0 = \sum_{n\in\mathbb{Z}_d} \alpha_{-n,-n}^n$, $\tilde{\alpha}_{r,s}^n = 0$ for all $n \neq 0$ and $r,\, s \in \mathbb{Z}_d$, and the rest of the entries in $\tilde{\alpha}$ coincide with their counterparts in α. In this process, the value of w_1 does not change but w_2 may increase. Hence, in our optimization, it suffices to study only those α such that $\alpha_{r,s}^n = 0$ for any $n \neq 0$ and $r,\, s \in \mathbb{Z}_d$ and, for the upper block $(\alpha_{r,s}^0)_{r,s} =: A$, $A \geq 0$ and $\mathrm{tr}[A] = 1$. From now on, we forget about the zero blocks and only concentrate on the upper block A. Let us denote the natural basis of $\mathbb{C}^d$ by $\{e_r\}_{r\in\mathbb{Z}_d}$ and define the Fourier operator $\mathcal{F} \in \mathcal{L}(\mathbb{C}^d)$ as in (5.1) with the basis $\{\varphi_j\}_j$ replaced by $\{e_n\}_n$. Fix the basis $\{f_n\}_n$, $f_n = \mathcal{F} e_n$, $n \in \mathbb{Z}_d$. We may define $w_1(A) = \langle e_0|Ae_0\rangle$, $w_2(A) = \langle f_0|Af_0\rangle$, and $w_0(A) = \min\{w_1(A), w_2(A)\}$. We have found that $W(\mathsf{A}, \mathrm{id})$ is the supremum of $w_0(A)$ over the positive trace-1 operators $A \in \mathcal{L}(\mathbb{C}^d)$.

This optimization task can still be further simplified: For any $A \in \mathcal{L}(\mathbb{C}^d)$, let us define $A^{\mathcal{F}} \in \mathcal{L}(\mathbb{C}^d)$ through

$$A^{\mathcal{F}} = \frac{1}{4} \sum_{k=1}^{4} \mathcal{F}^k A \mathcal{F}^{*k}.$$

The fixed points of the map $A \mapsto A^{\mathcal{F}}$ are exactly the Fourier-invariant operators B, i.e., $\mathcal{F}B = B\mathcal{F}$. One finds that $w_0(A^{\mathcal{F}}) = w_1(A^{\mathcal{F}}) = w_2(A^{\mathcal{F}}) = \frac{1}{2}(w_1(A) + w_2(A)) \geq w_0(A)$ for all positive $A \in \mathcal{L}(\mathbb{C}^d)$. Thus, our task is simply to optimize the linear functional $w_0(A) = \langle e_0|Ae_0\rangle$ over the set of positive trace-1 Fourier-invariant operators A on $\mathbb{C}^d$. The optimal value is reached at one of the extreme points of the set of such operators, and these extreme points coincide

with the projections onto the one-dimensional subspaces generated by the eigenvectors of $\mathcal{F}$. The Fourier operator has four eigenvalues, the fourth roots of 1 i^k, $k = 1, 2, 3, 4$, and the corresponding eigenprojections are

$$P_k = \frac{1}{4}\big(\mathbb{1} + (-i)^k\mathcal{F} + (-1)^k\mathcal{F}^2 + i^k\mathcal{F}^3\big), \quad k = 1, 2, 3, 4.$$

Hence, especially, for an extreme point A of the set of trace-1 positive Fourier-invariant operators, there is a unique $k \in \{1, 2, 3, 4\}$ such that $A = P_k A P_k$ and $w_0(A) = \langle P_k e_0 | A P_k e_0 \rangle$. One finds that $P_1 e_0 = P_3 e_0 = 0$ and $P_2 e_0 = \frac{1}{2}(e_0 - f_0)$ and $P_4 e_0 = \frac{1}{2}(e_0 + f_0)$ which means that w is maximized at a projection onto a one-dimensional subspace of the P_2- or P_4-eigenspace. It is obvious that, amongst such operators supported on the P_2-eigenspace, the optimal one is $A_- = |v_-\rangle\langle v_-|$ and, amongst the extreme points supported on the P_4-eigenspace, the highest value for w_0 is given by $A_+ = |v_+\rangle\langle v_+|$, where

$$v_\pm = \sqrt{\frac{\sqrt{d}}{2(\sqrt{d} \pm 1)}}(e_0 \pm f_0).$$

Simple check shows that $w_0(A_-) < w_0(A_+) = \frac{1}{2}(1 + 1/\sqrt{d})$. Thus, $W(\mathsf{A}, \mathrm{id}) = \frac{1}{2}(1 + 1/\sqrt{d})$. Again, for a decodable channel $\mathcal{V} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$, one has $(\mathsf{A}, \mathcal{V}) =_{\mathrm{post}} (\mathsf{A}, \mathrm{id})$, so that, according to Theorem 3:

Theorem 6. The robustness of incompatibility for any von Neumann observable $\mathsf{A} \in \mathbf{Obs}_d$ and any decodable channel $\mathcal{V} \in \mathbf{Ch}(\mathcal{H}, \mathcal{K})$ is

$$W(\mathsf{A}, \mathcal{V}) = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{d}}\Big).$$

The optimal positive trace-1 Fourier-invariant operator A_+ of the discussion preceding Theorem 6 gives rise to an operation $\mathcal{D}$ through Equation (5.22) which, in turn, defines an instrument Γ according to (5.18). It follows that

$$\Gamma_j(\rho) = \frac{\sqrt{d}}{2(\sqrt{d} + 1)}\Big(\frac{1}{\sqrt{d}}\mathbb{1} + \mathsf{A}_j\Big)\rho\Big(\frac{1}{\sqrt{d}}\mathbb{1} + \mathsf{A}_j\Big).$$

The optimal decomposition for the compatible pair $(\mathsf{M}, \mathcal{E}) = (\Gamma_{(1)}, \Gamma_{(2)})$ is given by

$$(\mathsf{M}, \mathcal{E}) = \frac{1}{2}\Big(1 + \frac{1}{\sqrt{d}}\Big)(\mathsf{A}, \mathrm{id}) + \frac{1}{2}\Big(1 - \frac{1}{\sqrt{d}}\Big)(\mathsf{B}, \mathcal{B}),$$

where

$$\mathsf{B} = -\frac{1}{d-1}\mathsf{A} + \frac{d}{d-1}\mathsf{T}, \qquad \mathcal{B} = -\frac{1}{d-1}\mathrm{id} + \frac{d}{d-1}\mathcal{E}_\mathsf{A},$$

where T is the trivial observable $\mathsf{T}_j = \frac{1}{d}\mathbb{1}$, $j \in \mathbb{Z}_d$, and $\mathcal{E}_\mathsf{A} \in \mathbf{Ch}_d^{W,2}$ is the Lüders channel $\rho \mapsto \sum_j \mathsf{A}_j \rho \mathsf{A}_j$ associated with A. Moreover, the optimal compatible pair $(\mathsf{M}, \mathcal{E})$ is given by

$$\mathsf{M} = \frac{\sqrt{d} + 2}{2(\sqrt{d} + 1)}\mathsf{A} + \frac{\sqrt{d}}{2(\sqrt{d} + 1)}\mathsf{T}, \qquad \mathcal{E} = \frac{\sqrt{d} + 2}{2(\sqrt{d} + 1)}\mathrm{id} + \frac{\sqrt{d}}{2(\sqrt{d} + 1)}\mathcal{E}_\mathsf{A}.$$

Similarly, for a decodable channel $\mathcal{V}$, the pair

$$\frac{1}{2}\Big(1 + \frac{1}{\sqrt{d}}\Big)(\mathsf{A}, \mathcal{V}) + \frac{1}{2}\Big(1 - \frac{1}{\sqrt{d}}\Big)(\mathsf{B}_\mathcal{V}, \mathcal{B}_\mathcal{V}) = \frac{\sqrt{d} + 2}{2(\sqrt{d} + 1)}(\mathsf{A}, \mathcal{V}) + \frac{\sqrt{d}}{2(\sqrt{d} + 1)}(\mathsf{T}, \mathcal{V} \circ \mathcal{E}_\mathsf{A})$$

is compatible, where

$$\mathsf{B}_\mathcal{V} = -\frac{1}{d-1}\mathsf{A} + \frac{d}{d-1}\mathsf{T}, \qquad \mathcal{B}_\mathcal{V} = -\frac{1}{d-1}\mathcal{V} + \frac{d}{d-1}\mathcal{V} \circ \mathcal{E}_\mathsf{A},$$

6. Conclusions

Given a convex subset L_0 of a real vector space, we have introduced measures of how well an element of the minimal affine subspace F containing L_0 resists immersion in L_0 under added noise. Especially, we have concentrated on the case where $L_0 \subset K \subset F$ and K is convex, in which case such measures can be defined relative to K. As a physical application, these robustness measures were studied in the case where K is the set of pairs of given quantum devices (observables or channels) and L_0 is the set of compatible pairs within K. In this context, we call such a measure as robustness of incompatibility.

Basic properties of the robustness measures have been investigated especially regarding monotonicity under certain compatibility-increasing partial orderings of device pairs. Lastly, values for the absolute robustness were calculated in three exemplary cases: a pair of Fourier-coupled rank-1 sharp observables, a pair of decodable channels (especially for unitary channels), and a pair consisting of a rank-1 sharp observable and a decodable channel.

However, we do not have a general method for how to calculate the robustness of incompatibility for a general pair of quantum devices; all our examples utilize symmetries in calculating the values of the robustness measure. Moreover, especially in the case of compatibility of observables and channels, it would, perhaps, be more natural to consider the other robustness function w instead of W; any noise in a measurement process affects both the registering branch and the state-change branch globally and is hence, typically, compatible. However, calculating the value for w in the example involving a von Neumann observable and a decodable channel leads to quite a complicated optimization problem. Moreover, calculating the robustness of incompatibility for the infinite-dimensional sharp Weyl-pair (position-momentum pair in $L^2(\mathbb{R})$) is a possible continuation of the analysis dealing with the finite-dimensional case of Section 5.1. For the time being, we only conjecture this infinite-dimensional pair to have the robustness $1/2$, i.e., this pair would be an example of a maximally incompatible pair according to the robustness measure.

References

[1] I. Bengtsson and K. Życzkowski, "Geometry of Quantum States, an Introduction to Quantum Entanglement" (Cambridge University Press, New York, 2006)

[2] N. Brunner, M.T. Quintino, and T. Vértesi, Joint measurability, Einstein-Podolsky-Rosen steering, and Bell nonlocality, Phys. Rev. Lett. 113, 160402 (2014)

[3] F. Buscemi, M. Keyl, G.M. D'Ariano, P. Perinotti, and R.F. Werner, Clean positive operator valued measures, *J. Math. Phys.* **46**, 082109 (2005)

[4] P. Busch, T. Heinosaari, J. Schultz, and N. Stevens, Comparing the degrees of incompatibility inherent in probabilistic physical theories, *EPL*, **103**, 10002 (2013)

[5] C. Carmeli, T. Heinosaari, and A. Toigo, Informationally complete joint measurements on finite quantum systems, *Phys. Rev. A* **85**, 012109 (2012)

[6] G. Cassinelli, E. De Vito, and A. Levrero, On the decompositions of the quantum state, *J. Math. Anal. Appl.* **210**, 472-483 (1997).

[7] T. Eggeling and R. F. Werner, Separability properties of tripartite states with $U \otimes U \otimes U$ symmetry, *Phys. Rev. A* **63**, 042111 (2001)

[8] O. Gühne, T. Moroder, and R. Uola, Joint Measurability of Generalized Measurements Implies Classicality, *Phys. Rev. Lett.* **113**, 160403 (2014)

[9] E. Haapasalo, T. Heinosaari, and J.-P. Pellonpää, When do pieces determine the whole? Extreme marginals of a completely positive map, *Rev. Math. Phys.* **26**, 1450002 (2014)

[10] E. Haapasalo, M. Sedlák, and M. Ziman, Distance to boundary and minimum-error discrimination, *Phys. Rev. A* **89**, 062303 (2014)

[11] T. Heinonen, Optimal measurements in quantum mechanics, *Phys. Lett. A* **346**, 77-86 (2005)

[12] T. Heinosaari, J. Kiukas, and D. Reitzner, Noise robustness of the incompatibility of quantum measurements, arXiv:1501.04554v1 [quant-ph] (2015)

[13] T. Heinosaari, T. Miyadera, and D. Reitzner, Strongly incompatible quantum devices, *Found. Phys.* **44**, 34-57 (2014)

[14] T. Heinosaari, J. Schultz, A. Toigo, and M. Ziman, Maximally incompatible quantum observables, *Phys. Lett. A* **378**, 1695-1699 (2014)

[15] A.S. Holevo, Radon-Nikodym derivatives of quantum instruments, *J. Math. Phys.* **39**, 1373-1387 (1998);

[16] A.S. Holevo, R.F. Werner, Evaluating capacities of bosonic Gaussian channels, *Phys. Rev. A* **63**, 032312 (2001)

[17] A. Jenčová and D. Petz, Sufficiency in quantum statistical inference, *Comm. Math. Phys.* **263**, 259-276 (2006)

[18] A. Jenčová and S. Pulmannová, How sharp are PV measures?, *Rep. Math. Phys.* **59**, 257-266 (2007)

[19] A. Jenčová, S. Pulmannová, and E. Vinceková, Sharp and fuzzy observables on effect algebras, *Int. J. Theor. Phys.* **47**, 125-148 (2008)

[20] P. Lahti, Coexistence and joint measurability in quantum mechanics, *Int. J. Theor. Phys.* **42**, 893-906 (2003)

[21] H. Martens and W.M. de Muynck, Nonideal quantum measurements, *Found. Phys.* **20**, 255 (1990)

[22] M.F. Pusey, Verifying the quantumness of a channel with an untrusted device, *J. Opt. Soc. Am. B* **32**, A56 (2015)

[23] P. Skrzypczyk, M. Navascués, and D. Cavalcanti, Quantifying Einstein-Podolsky-Rosen steering, *Phys. Rev. Lett.* **112**, 180404 (2014)

[24] G. Vidal and R. Tarrach, Robustness of entanglement, *Phys. Rev. A* **59**, 141-155 (1998)

E-mail address: ethaap@utu.fi

Turku Centre for Quantum Physics, Department of Physics and Astronomy, University of Turku, FI-20014 Turku, Finland